ELSEVIER

Coordination Chemistry Reviews
200–202 (2000) 247–348

COORDINATION CHEMISTRY REVIEWS

www.elsevier.com/locate/ccr

The distribution of elements in cells

R.J.P. Williams [a,*], J.J.R. Fraústo da Silva [b]

[a] *Inorganic Chemistry Laboratory, Oxford University, South Parks Road, Oxford OX1 3QR, UK*
[b] *Instituto Superior Técnico, Universidade Técnica, Lisbon, Portugal*

Received 25 August 1999; accepted 27 February 2000

Contents

* Corresponding author. Tel.: +44-1865-272600; fax: +44-1865-272690.
E-mail address: bob.williams@chem.ox.ac.uk (R.J.P. Williams).

PII: S0010-8545(00)00324-6

Abstract

The selective uptake of elements by proteins can be based on several routes: (a) the equilibrium binding of different elements by different protein ligands using charge, size, electron affinity and stereochemical preferences; (2) kinetic insertion of an element into such a coordination site of a protein; (3) removal of the element to a special compartment by pumping followed by (1) or (2). A cellular system also limits the amount of each type of metal-binding apoprotein by genetic regulation of its symbiosis with element uptake. Such a limitation generates much greater selectivity. Finally we consider how the observed selection of elements by proteins has changed in evolution through changes of availability of elements and their combinations in the environment [1].

Keywords: Element distribution; Equilibria; Insertion; Membrane pumping; DNA expression; Feed-back controls

1. General introduction

1.1. Cells and genetic instructions

In order to understand the distribution of elements, especially the so-called inorganic elements, in cells [1–3][1] it is necessary to give an outline description of a cell. While a bacterial cell has but an inner cytoplasm contained by an outer membrane and wall, or in some cases a double membrane and a wall with a space between the membranes called the periplasm, a eukaryotic cell contains a series of compartments — extracellular space (periplasm), cytoplasm, nucleus, vesicles and reticula, and organelles each contained by a membrane. Each compartment has its own environment, pH, salts etc., and each has its own complement of organic compounds and inorganic elements. The positioning of all this material in the cell depends upon binding sites and transfer processes to the binding sites. Both are also intimately associated with proteins and hence their production. Thus understanding the distribution of elements requires both knowledge of the handling of elements and of the synthesis, under the control of the genetic code, and purposeful handling

[1] These three books are a trilogy. In the first complex ion chemistry is described and then this knowledge is used in a comparative study of *isolated* biological molecules containing special elements not in organic biological chemistry. This is the usual approach to biological inorganic chemistry. In the second volume interest lies in the evolution of the integrated use of the elements in compounds in organisms treated as systems more than as individual molecules. The third book is an examination of the underlying principles of species differences of organisms, uncovered in chemical detail in the first two books, linking these differences to the principles of speciation in chemistry and to the complexity of properties of all substances in an environment. The complication of organisms lies in the fact that they are complex systems in flow, which through the help of the continuity of a code, can evolve with the changing environment.

of many proteins. The first section below gives a very brief general introduction to the *regulation* of protein production, directly related to genetic (DNA) structure, and to the *control* of the activities of the proteins once produced.

In this article we shall use the word gene rather broadly to refer to a functional length of DNA which at its simplest (i) translates ultimately into a single protein product from a reading frame of the DNA, or (ii) regulates production, in that the condition of it determines the possibility of expressing a protein, where this length of DNA is called a transcription factor binding region, or (iii) it is a binding, promoter, region for the machinery which reads the DNA, a polymerase (enzyme) site, Fig. 1. There are also regions of the code which appear to be meaningless, called introns, but they may well have structural significance, allowing bending of DNA for example. Only 5% of human DNA reads directly as protein. Using this sense of the word 'gene', a gene is not necessarily very closely related to a particular characteristic of an organism, though it may be directly related to production of some required protein involved in helping to generate the characteristic [4,5].

We take it that in a cell management of protein activity and the production of all proteins, which control the concentrations of all small chemicals including inorganic elements, and RNA is *regulated* eventually at the genetic, DNA level, see below. Control of concentrations is also dependent upon the supply of both energy and material to protein machinery, for capture, synthesis and pumping, and eventually is dependent on the environment. For example in the production of pyruvate from glucose by the glycolytic pathway the controls over the rate of production are due to: (1) the presence of a series of enzymes, carriers and pumps (see Fig. 3) and their inhibition by small molecules (*control*); (2) the production of these enzymes, carriers and pumps synthesised by translation at the RNA level; and (3) the transcription of genes as messengerRNA at the DNA level. We shall combine (2) and (3) under the term *regulation*, which includes decrease or increase in management of protein production. Finally in advanced cells there is (4) the need to control the supply of glucose to a cell environment through the use of organs, controlled by hormones, which release the sugar into circulating fluids. The handling of an element in free ionic form or in a more complicated molecule is dependent upon exactly the same factors.

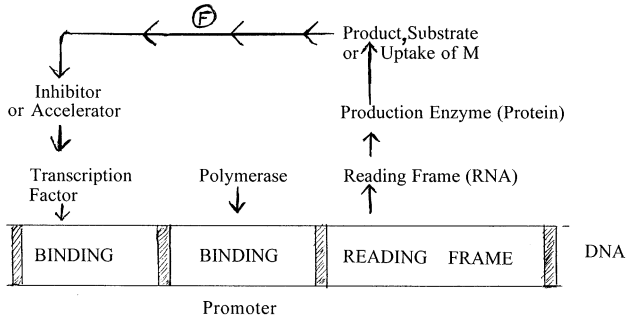


Fig. 1. An outline description of the important functional features of DNA in the expression of proteins and the effect of products in feedback (F). The element, M, acts with or without a protein on DNA.

1.2. Control

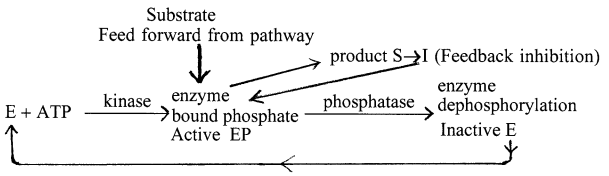
Taking control first we write for the rate of a very simple reaction of a substrate, $S \xrightarrow{E}$ Product, catalysed by an enzyme, E, where *equilibrium formation* of $|ES|$ is rate limiting, Eq. (1).

$$\frac{d|S|}{dt} = \frac{k_1 K_1 |E_T|}{1 + K_1|S| + K_I|I|} \quad (1)$$

Here $[E_T]$ is the total enzyme concentration, and I an inhibitor which can be a product of the pathways of synthesis or degradation of S when I is a feedback inhibitor. The constants K_I and K_1 are equilibrium binding constants for inhibitor and substrate, respectively, and k_1 is a rate constant. Of course there are many more complicated rate equations for enzyme reactions and their controls including accelerators of rates. Additionally the supply of S may be controlled by environmental levels of precursors and their uptake pumps and exchangers as well as by the products from other pathways, which may increase or decrease the rate. Clearly while S can be a conventional organic molecule it can also be a metal ion undergoing redox or any other transformation such as incorporation in a protein. Feedback control can also apply to the pumps and exchangers, where pumping is across membranes, when S, the substrate of the pump (exchanger), may self-inhibit reaction once [S] in the cell becomes high. It is easy to see that each pathway, enzymes + pumps (exchangers), needs to be controlled but it is also necessary for cooperative cellular activity to have rates of transformation in one pathway linked to rates in others — synthesis and degradation and uptake and rejection, for example, must not be independent. In other words a cell has to be of fixed or almost fixed composition in a steady state after development while maintaining synthesis and degradation and uptake and loss. Therefore the machinery in a given pathway has to be inter-connected by inter-pathway messenger control systems as well as by an internal control in its own pathway. As an example it is essential that all the pathways involving condensation, that is all the major polymerisation reactions, for example protein and RNA production of a cell, run together at fixed relative rates. We know that the reactions, which produce the combination of ribosomal RNA, rRNA, and ribosomal proteins, do not give rise to excess of either polymer. This can only be managed if the organic compound pathways for their syntheses and degradations are based on common communicating networks of organic material transfer, which we know is particularly due to the co-enzymes which carry and distribute H, C, N, O elements in fragments of conventional substrates to different enzyme pathways. The co-enzymes may also act as inhibitors or stimulators of steps in the pathways. Of course substrates themselves can act as feedback inhibitors. Equally we know that the channels of uptake and the pumps for rejection of an ion such as calcium are balanced to give a fixed homeostatic level of free calcium by feedback of $[Ca^{2+}]$ to inhibit or accelerate transfer of the ion.

Apart from the supply of H, C, N, and O all the organic compound pathways also need energy to drive many reactions, e.g. condensations. The energy derives from NTP, nucleotide triphosphates, especially ATP, and thus ATP is entirely

suited as one communicating control factor between many pathways. A common control by ATP is in fact based on phosphorylation/dephosphorylation reactions of proteins — a kinetic scheme of phosphate binding and release is as follows:



If all the control kinases have a very similar binding constant for ATP then all kinetic pathways in a cell can run under parallel control. They also have individual feedback off-switches to their enzymes from particular individual products, I, of given pathways, as above, so that they do not run at the same rates and accumulation of unused products is minimal. The common use of ATP as an energy source is not confined just to the various pathways but applies to pumps and also to controls of the translation and transcription of the cell, which we treat in more detail in Section 1.3. We shall consider the balanced introduction of the chemical elements other than H, C, N and O into the different pathways later.

We see that a cell manages to produce a coherent set of molecules, proteins, nucleotides, saccharides and fats (and as we see later inorganic elements) largely through the use of a few molecules which act both as carriers of material and energy and as feedback controls on all the steps associated with the pathway enzymes and pumps and on their production through DNA expression. As stated these molecules are mainly the common *mobile* co-enzymes. They act as the units of integration throughout the cell and it may be helpful to look upon one of their features as being similar to electrons in an integrated electronic circuit. Some metal ions act in a parallel manner as we shall see. The selectivity of molecular recognition allows the simultaneous use in space and time of several current carriers and a multiplicity of circuits unlike electronic systems. When substances in different paths are very similar, e.g. SO_4^{2-}, MoO_4^{2-} and HPO_4^{-}, or Mg^{2+}, Ca^{2+} and Mn^{2+}, we also have to see how the surface of a protein can distinguish and select one from another, see below.

We can now understand in essence how a simple cell can take in and utilise chemical elements in a coordinated and balanced way utilising a given set of catalysts, pumps (and exchangers) and carriers (see Fig. 3). For a detailed account of *control* over uptake and substrate flow the reader should refer to modern books on biochemistry and flux analysis, see references [6,7]. We must also keep in mind that the catalysts, carriers and pumps, all of which are proteins, are themselves produced at regulated rates by transcription (DNA to RNA) and then translation (RNA to proteins) and that there are different factors managing the productions of different proteins.

1.3. Regulation: constitutive and induced protein production using genes [4,5,8]

For operational purposes it is useful for us to separate genes into classes, constitutive and induced, no matter if in practice the distinction we make is not so clear cut.

A *constitutive* gene is responsible for the regulated production of a protein, which is essential for the well being of the cell in all normal circumstances. For example some proteins are produced in a controlled concerted fashion with the cell cycle but for most the reading of the constitutive gene must be switched off in other ways so that excess protein is not generated. The usual regulation process is one in which a given substance, a protein or a small molecule generated directly or indirectly by the gene and which must first be produced to a certain reasonable concentration, builds up in concentration so that eventually it becomes a feedback inhibitor of production by binding to a transcription factor (see Fig. 1 and Eq. (1)). This feedback, called *end-product repression*, is necessary to ensure that energy is not wasted and a given protein is not produced excessively. Frequently here the feedback inhibition is associated with a relatively weak binding constant to the transcription factor and hence a weak restrictive effect on DNA by the active participant in repression, that is the protein or substrate bound to a transcription factor, since an optimal reasonable free concentration of the protein or substrate must be synthesised and kept at a given level for the running of the cell. Notice that a transcription factor must be constitutively produced often by cell cycle limitation.

An *inducible* gene is responsible for looking after a cell in unusual circumstances so that expression of its protein is not required normally. The gene is repressed by some particular transcription factor protein, which binds to it strongly. However under stress of some kind the effect of the stress on the cell leads to a signal which removes repression and leads the gene to express the inducible system so as to relieve the stress. The stress itself or the stressed condition of the cell, its adjusted metabolism, informs the genome that the induced system should be switched on. In the absence of the stress the system is not noticeably expressed. Clearly the switching on of an inducible system may well have to be very sensitive to the inducer (or poison) so that a high-binding constant of the inducer to the protein, transcription factor, which represses the expression of the DNA is required in the feed-forward activation of transcription and translation. Note that since the repression is due to binding of a repressor there will always be some production of a protein since binding constants are not infinitely large. While many genes are repressed under particular circumstances there are also positive regulatory genes which enhance gene expression.

Examples help to see the way the genes and their regulation work. In *Escherichia coli* the glycolytic pathway enzymes are produced constitutively in tune with the cell cycle [8]. Amino acid synthesis is also essential for the survival of bacterial cells hence amino acid synthesis is also constitutive, however it suffers end-product repression. Taking the examples of histidine and tryptophan: the synthesis of these two amino acids has to fit requirements for the production of say 1000 proteins. However, there must not be excess of either amino acid, or of any other over

Table 1
Levels of required metal ions which can be poisonous in the cytoplasm[a]

Poisonous concentration (M)	10^{-1}	10^{-3}	10^{-5}	10^{-9}	10^{-10}
Metal ion	Na^+, Mg^{2+}	Ca^{2+}	Mn^{2+}, Fe^{2+}	Zn^{2+}, Cu^{2+}	Ni^{2+}, Co^{2+}

[a] N.B. Many metal ions are poisonous at almost any concentration, e.g. Hg^{2+}, Pb^{2+}, Ag^+.

requirement, hence both histidine and tryptophan themselves as they build up in concentration begin to bind to transcription factors which then stop further synthesis of these very amino acids. This is *negative feedback*, or end-product repression, and requires a relatively weak-binding constant of $< 10^6$ so that free amino acid levels never fall below this optimal value.

The situation is made somewhat more complex in that when supply of an essential substance runs low it pays the cell to increase the production of the proteins for uptake of that substance so that the transcription of relevant genes may be *enhanced* by binding of a different transcription factor not bound to the substance to be pumped but to some later product of metabolism. This is an example of *positive feed-forward*.

Consider next a man-made drug that is not readily metabolised and the metabolic products of which are valueless. After suitable mutations of genes pre-existing the ability to remove the drug, there is generated a gene for destruction of the drug. Now this gene product has to be produced in effective amounts only in the presence of the drug although the gene is now a part of the inherited genome. Provided that the drug binds to a transcription factor with a high affinity, and that only this combination binds DNA, there is only effective production of the gene product when the drug is present. The drug induces the enzyme by *positive feed-forward*. In bacteria such inducible genes are often on a separate small cyclic piece of DNA termed a plasmid. Note that it may not be concluded that absolutely no copies of an 'inducible' protein are produced in the absence of inducer. All of this account of gene regulation is to be found in more detail in references [4,5].

Now we are interested in this article in the handling of the chemical elements by cells and hence we need to know which elements are looked after constitutively and which by induced systems. In many cases it is found that a cell has to be flexible in that it has to adjust gene activity so as to control levels of proteins and associated metal ions and non-metal elements in cells in close harmony opposite a variable environment. This means that both repression and acceleration, feedback and feed-forward systems, are required over the element uptake/rejection as well as over the protein concentrations. To meet this requirement uptake and rejection systems for elements co-exist in the endeavour to maintain homeostasis. Individual elements must therefore be associated with selective binding to transcription factor proteins. We want to put this discussion on as quantitative basis as possible but we have to remember that some required elements are poisonous in the cytoplasm at very low levels, Table 1. We expect these elements, e.g. Hg^{2+}, to be selectively handled by strong induced binding while required elements such as K^+ or Mg^{2+} will be handled by weak constitutive binding.

1.4. Constitutive and induced genes for elements

We consider in the first instance bacterial cells, in particular *E. coli* [8]. An element, for example a metal ion, M_i, is required at selected level in cells but excess is damaging. Hence there has to be adequate production of constitutive uptake systems, inward channels of pumps or exchangers, especially if the element is in short supply, and means of switching off the production when there is an adequate number. It is also necessary to pump the metal ion outwards if the element can gain access to the cell to a larger degree than required and outward pumps must also be produced in some regulated mode. Rate of production of the pump proteins can be inhibited or enhanced as discussed above so that we can write rate equations for pump (P_i) production for a given internal concentration of the metal ion, [M_i].

$$\frac{d[P_i]}{dt} = \frac{k_2|X_T|}{1 + K_2|M_i|} \tag{2}$$

where X is the transcription factor (TF), for pump production, which binds M_i, and can be bound to DNA and [X_T] is its total free concentration. The sign of k_2 can indicate outward or inward pumping. The constitutive rate constant for production of the pump to maintain growth is k_2 but the production is shut down when [M_i] is large. K_2 is the binding constant of M_i to the transcription factor, where M_iX switches off production of P. It is assumed that M_iX fails to bind to DNA, but X does. There needs to be constitutive production of this transcription factor which may also depend upon the concentration of M_i in a similar way. If the transcription factor, either bound to M_i or free, is always bound to DNA stoichiometrically and assuming the metal free form activates in one conformation but deactivates when metal bound and in a different conformation the equation for pump production is

$$\frac{d[P_i]}{dt} = \frac{k_2}{1 + K_2[M_i]} \tag{3}$$

It can be the case that on entering the cell the element M_i is bound first to a carrier C_i and eventually placed in a functional unit, E_i, for example an enzyme, while also interacting with a transcription factor, TF. In all cases the expression rate of the protein production during the growth cycle takes the same form as above so that free metal ion is allowed to reach a certain level but is then restricted by binding constants. Thus the proteins are produced commensurately with the element concentration so that there will be related free M_i and free apoproteins. The production of any of the proteins is eventually stopped by feedback inhibition but in some cases it can be enhanced over basic constitutive rates by other transcription factors, Y, which bind the element when the expression is more complex:

$$\frac{dP_i}{dt} = k_2\left\{\frac{|X_T|}{1 + K_2|M_i|} + \frac{K_3|M_i||Y_T|}{1 + K_3|M_i|^2}\right\} \tag{4}$$

where the second term is an effective 'induction' constant and K_3 is the binding constant of M_i to the appropriate transcription factor Y for enhancement. Repression not due to a single binding factor, M_iX, but to a factor M_iY, is described using the same equation with the reverse sign of the second term.

At this point it is worth observing that if all the above proteins, transcription factors, carriers, pumps and enzymes, bind to an element in exchange equilibrium then for a given element there has to be a similar binding constant to each of the proteins. For strong binding, the implication is that in the steady state, all the proteins are overwhelmingly bound to M_i and there can then be very little free M_i and very little free protein, TF, P, E, or C, when the cell is running optimally. The sum of all of these proteins is effectively stoichiometrically related to the bound M_i. Clearly this will be the case for the more poisonous metal ions of Table 1. We shall see how this can be achieved in the case of mercury ions by having all its genes under the control of one mercury binding transcription factor.

Now the rate of induction of a pump, carrier, enzyme or buffer protein takes a similar form to the second term of Eq. (5) and is dependent on M_iZ_T:

$$\frac{dP_i}{dt} = \frac{k_4K_4|M_i||Z_T|}{1 + K_4|M_i|}$$

$$\frac{dP_i}{dt} = \frac{k_4K_4|M_i||Z_T|}{1 + K_4|M_i|} \tag{5}$$

where pumping may be outward. There is no constitutive term and no activity until M_i is present externally to give some internal $[M_i]$. K_4 is the binding constant of M_i to the transcription factor Z_T for an outwardly directed pump if M_i is to be removed, k_4 gives the optimal rate. For induced inwardly directed pumps the rate of production is of the same form. Provided inwardly directed pumping has a higher affinity but a lower rate than outwardly directed pumping then a steady state $|M_i|$ is achieved. The pump production can be regulated by the action of repression due to a concentration of M_i above that required. Metal ions which have no function but are poisonous must be rejected by systems of proteins with high binding constants to such metal ions.

The genes for these proteins may well be on a single *operon* so that the DNA has a series of coded messages

X-Y-Z-polymerase-P-C-E

while production of the transcription factors X, Y and Z is due to genes elsewhere in the DNA. P, C and E stand for required protein pumps, carriers and enzymes for a given M_i. Production of E may well depend on pumps, P, and carriers, C, as well as transcription factors, X, Y and Z due to forward feed from M_i.

1.5. Summary of control and regulation due to equilibria

All the above Eqs. (1)–(5) use equilibrium binding constants, K, which may control activity or regulation of synthesis of a protein. As stated all K values must

be almost equal for the handling of one substance M_i in a bound state opposite free M_i to ensure equal activity. It is then very important to know which processes can be safely described as being at equilibrium. This returns us to the problem of describing the modes of incorporation or rejection of given elements, since some incorporations, as we shall see, are under kinetic control. Before doing so we give a table of elements the incorporation of which is constitutive, rather than induced, in *E. coli* and maybe generally so (Table 2).

While we see that for any one metal ion the binding constants to its selected proteins must all be rather similar for all reactions which equilibrate with a given free M_i, it is also imperative that under the conditions that exist in the cell, no other metal ion, M_2, binds to these same sites. One of the severe difficulties of control and regulation that is of necessity at equilibrium or close to it, is that such selectivity is required. We shall indicate how this is possible, through the application of 'conditional' (or effective) equilibrium constants, even when a cell has to handle at least ten similar elements simultaneously, all of those listed under constitutive in Table 2 in fact. A final implication of this tight feedback control over equilibrium concentrations of free ions is that to a first approximation the free ion concentration of each metal ion is the inverse of its binding constant. This follows from the stoichiometric uptake of each metal ion opposite the sum of all ligands designed to bind it and the nature of stability constants

$$T_L = |L| + |ML|$$

$$T_M = |M| + |ML| \text{ therefore } |L| = |M| \text{ since } T_M \simeq T_L$$

$$K = \frac{|ML|}{|M||L|} = \frac{|ML|}{|M|^2} \simeq \frac{T_{ML}}{|M|^2} \tag{6}$$

Let us suppose that the metal protein is 10^{-4} M in the cell then

$$K \times 10^4 = 1/|M|^2$$

The free $|M|$ is inversely proportional to K and we may expect values for K and $|M|$ to be approximately as in Table 3 and Fig. 2. Of course these values only give a rough impression.

Table 2
Constitutive and induced systems for element incorporation or rejection in many cells [8][a]

Constitutive incorporation or partial rejection (Essential elements)	H, C, N, O, for essential organic metabolism Na, Mg, (Si), P, S, Cl, K, Ca, (V) (Cr) Mn, Fe, Co, (Ni), (Cu) Zn Mo, Se, I
Induced rejection (poisons)	(Al), As, Cd, Ba, most heavy elements after molybdenum in the periodic table N.B. Many foreign organic molecules, drugs

[a] Rejection can be by metabolism when an enzyme not a pump is induced.

Table 3
Free metal ion concentrations in cells

Metal ion	Ligand (M)	K	Free metal ion (M)	
	(Total maximum)		Calculated	Observed [1]
Mg^{2+}	ATP (10^{-3})	10^{4}	$10^{-3.5}$	$10^{-3.1}$
Fe^{2+}	Protein (10^{-5})	10^{8}	$10^{-6.5}$	10^{-8}
Zn^{2+}	Protein (10^{-5})	10^{12}	$10^{-8.5}$	10^{-11}
Cu^{+}	Protein (10^{-5})	10^{15}	10^{-15}	$<10^{-12}$

In this article we are overwhelmingly concerned with the handling of inorganic elements, M, not C/H/N/O compounds and therefore while we have given an introductory account of control and regulation based on such organic molecules as glucose and histidine in what follows we shall refer very largely to elements in simpler compounds in organisms, e.g. Fe^{2+} and SO_4^{2-}.

1.6. A note on more complicated management of flow (flux) [9]

It is now generally agreed that the *control* of metabolite flow through a pathway does not reside in the feedback inhibition of single enzymes. Control rests upon a multitude of feedback interventions and also upon control of supply, which includes pumps and channels. The treatment of flow is then best understood in terms of overall system analysis. It is often a useful approximation to focus upon the metabolism of one particular substance, for example glucose, and its pathways for illustrative purposes of how control works but even here it is found that especially in higher organisms the overall management of flow rests with such

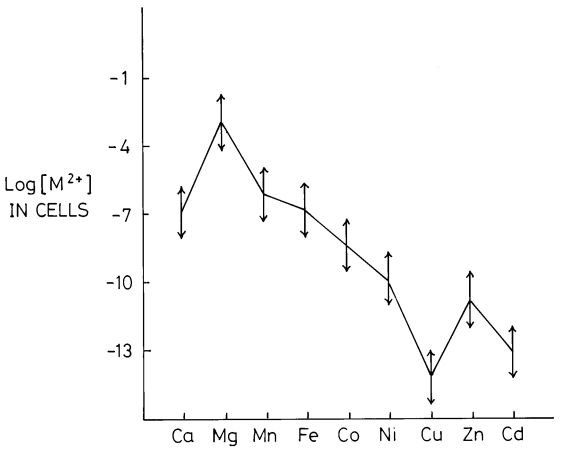


Fig. 2. The logarithm of the approximate concentration of *free* metal ions of different kinds in the cytoplasm of cells generally.

factors as hormone release, which affect *regulation* (enzyme, pump and carrier production) as well as multiple feedback control of metabolites. The situation is not different for the pathways of elements, e.g. metal ions, which are usually handled in a series of steps. At each rate-limiting step there can be feedback control, e.g. at a pump or at an insertion. However a step which equilibrates cannot be controlled, except by: (a) the total amount of material in the system; or (b) the restriction to a compartment. In the uptake or rejection of an element it is also observed that there is *regulation* of the production of enzymes, pumps, exchangers or carriers. It is also the case that the system for handling one element is interactive with the handling of more than one other element so that the whole organism activity is an interactive network. In such circumstances as in the discussion of metabolic pathways it is necessary to simplify in order to show how different elements are managed separately before showing the interactions between them. At the same time we must observe that flow in a network demands energy. The simple consideration of an electrical circuit which is either in constant flow or regularly pulsed gives an impression of communication networks in biological cells. These networks are more complicated since there are many running at once contacting one another only here and there. This is not easily achieved and requires selective recognition in each circuit of a given current carrier. Note again that electrical circuits have but one current carrier-the electron, and are useful for comparison but not detailed insight into the functioning of a cell.

1.7. DNA: mutations and knock-out experiments

We include this section in order to show how difficult it is to be sure of the functional value of a protein. The presence of the code allows experiments to be made on the requirement for a given protein and/or an individual amino acid and it also allowed evolution to occur by changing, inserting, or deleting bases in the DNA sequence so as to modify proteins by changing their amino acid composition. The procedure of mutational analysis can help the understanding of action of individual amino acid residues in maintaining activity. Following DNA/RNA changes in different organisms also helps us to understand evolution. In the case of metallo-proteins it is particularly interesting to evaluate the functions of amino acids in the coordination sphere or near neighbours to it. Particularly extensive investigations have been carried out on the blue copper proteins when the effect on redox potentials and electron transfer rates have been studied [10]. There is some controversy but generally the work has supported the idea that the protein generates a *constrained site* (variously also called an entatic state or sometimes an induced rack condition) for optimal action of the metal ion, see sections on specific enzymes.

A different approach is to remove (knock-out) a particular gene, from the genome and to study the effect on a whole organism. Here the aim is to discover the function of the protein codified by this gene. The results sometimes show that removal of a single gene has very little effect even when followed over more than one generation, say of breeding mice. Examples are the S-100 calcium binding

proteins and the metallo-thioneins. The suggestion is that the genome as a whole can generate alternative products to salvage such situations for otherwise the protein would appear to have no function. In other cases the knock-out experiment leads to so many faults in embryonic stages that the observations can only be said to be confusing but the gene and the product may have multiple roles. In a few cases such as the removal of the gene for the calcium-binding protein, parvalbumin, the result is definitive [11]. Here the knock-out of parvalbumin slows the recovery stage of muscle relaxation. Therefore parvalbumin has been proved to be a muscle relaxation factor as was suggested by many early experiments.

We return to the discussion of the genome in evolution and the effects of environmental stress on its changes in Section 9.

1.8. Thermodynamic and kinetic controls

All the analysis so far is based upon equilibria between the substrate or metal ion and the protein to which it binds. The reactions are in one compartment and include binding to the surfaces of pumps and DNA in the containing membrane of that compartment. In cellular systems there are two other possible selection steps we have to consider: (1) the passage through a pump to isolation in a compartment; and (2) the kinetic uptake, one-way combination of the substrate or metal ion with a protein. Both processes are irreversible. In the case of the pump the irreversible step is the transfer of M_i from one side of the membrane to the other using external

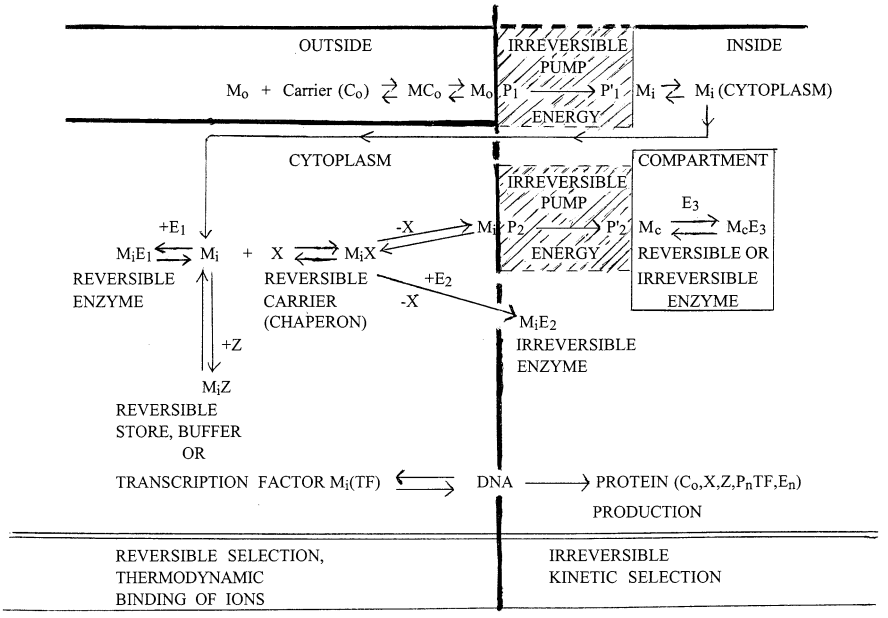


Fig. 3. A schematic representation of the uptake modes for different elements, M_i, into compartments and then their selective incorporation by thermodynamic (equilibrium) or kinetic traps.

energy often from ATP. The irreversibility is then only in the membrane where binding to a centre is altered by a conformation change with dissipation of heat. The binding to the site from either side in their different conformation states is reversible, see Fig. 3.

The second case of effective one way reaction can come about by simple covalent bond formation which may or may not require energy. Capture of elements such as carbon from CO_2 or nitrogen from N_2 requires energy, again from hydrolysis of ATP. There is another possibility that of insertion into a kinetic trap, for example of Mg^{2+} into chlorin, which may or may not require energy. Finally an element can be trapped by the strength of its binding brought about by the folding step of metallo-protein formation.

We can represent all these steps in one diagram showing how the final distribution of an element can be reversible in part and irreversible in part. In many ways the reversible part is the more interesting since it allows the element shown as M_i in Fig. 3 to connect to a great variety of proteins including those managing its distribution. The reversible steps can be treated by equilibrium constants which are easier to understand than kinetic steps and barriers. We shall proceed by indicating the nature of the different paths for different elements so that the relative importance of reversible equilibration and irreversible steps can be seen.

2. General discussion of observed distribution of elements in cells

2.1. The distribution of elements in organisms

A cell contains some twenty different chemical elements [1–3], Table 2. Critical for the cell's function is the way in which the elements are distributed in it. Here distribution includes the chemical forms and combinations of the element as well as the positioning of the chemical species in different parts of cellular space. That these distributions are fixed features of organisms implies that the mechanism of distribution is highly organised (see Fig. 3). As stated the coordination is managed centrally (*regulated*) in cells by the genome while the genome itself is in contact with gene products, including incorporated ions, and their distributions. Additional organisation is *controlled* at the level of back-interaction of element concentrations with pumps, channels, carriers, enzymes and exchangers which move the element through membranes and cells and then bind them more permanently. It is therefore intense feedback interactions that maintain the distribution and ensure homeostasis of each element. While some atoms of elements are in statically fixed positions others flow in a sequence in the organism. All 20 elements are handled quite specifically so that their physical and chemical positioning is free from contamination by other elements. We would like to know how the individual elements are handled so as to maintain the overall dynamic homeostasis. Clearly the organisation required and achieved is not based on a repetitive lattice as in crystalline chemicals but on an inhomogeneous yet particular pattern of flow [3]. For more complicated organisms controlled extracellular fluids must also be managed.

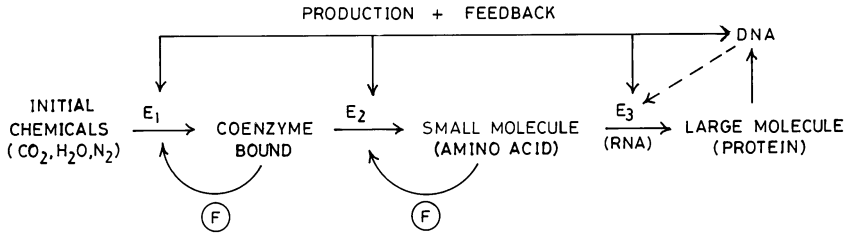


Fig. 4. A very much simplified scheme for the uptake of the major elements, C, H, N and O through covalent intermediates. E is an enzyme, (F) is feedback.

Provided that we can understand the way in which elements are distributed in cells under different environmental circumstances we can approach the way in which distribution has changed during the four billion years of evolution, Section 8. To this end it is necessary to also appreciate the changes in element speciation in the environment with time since availability is a critical factor for access of an element to an organism [2,3].

Before proceeding note again that distribution of an element has two features: (a) incorporation in selected compounds; and (b) the disposition of this compound in the different parts of particular compartments of a cell and of an organism. We wish to uncover not only the compartmental whereabouts of elements but also how they have become positioned both in particular chemicals and in particular places within selected phases.

2.2. Major chemical elements in cells

The major chemical elements, C, H, O and N, form all the main parts of the cell. The materials themselves are organic molecules which are kinetically stable and include polymers such as proteins, fatty acids, polynucleotides and polysaccharides. These different molecules are required in relatively fixed amounts but since the composition of each in terms of element, C, H, N, O, content is different, there has to be a way of distributing the elements commensurately into the major synthetic pathways. An example, given earlier, is the almost exact synthesis of certain proteins and rRNA to make ribosomes with no excess of any component. The distribution of elements utilises initially controlled catalytic pathways, e.g. glucose, amino acid, or nucleotide synthesis. These pathways take up or release elements from or to carrier mobile co-enzymes, which themselves separately handle C, H, N or O in useful units. To ensure appropriate incorporation then demands that the elements, initially in the form of small molecules such as H_2O, CO_2, and N_2, are incorporated into suitable transferable units, co-enzymes such as H in NADH, CH_3CO– in acetyl Co-A, CH_3– in methyl methionine CoA and NH_2– in glutamine. The initial small molecules themselves (H_2O, CO_2 and N_2) pass through membranes readily, so that controlled pumps for them are not essential. Enzyme pathways which recognise these small molecules and intermediate fragments control the subsequent build-up into larger and larger molecules. The production of some of

these enzymes of the pathways is *constitutive*, although it may also be *enhanced*, while other enzymes for more specialised reactions are *induced*, so that much management is at the level of DNA regulation, Fig. 4. Of course the enzymes only operate if there is an adequate supply of the initial small molecules. Each pathway is feedback managed at the enzymes often by the products of later steps in it so that there is no excessive build up of one or another small or larger molecule, Fig. 4. We note immediately that energy is utilised throughout the uptake, distribution and incorporation so that none of it is under equilibrium conditions and the whole system is completely irreversible. The energy supply is critical and is distributed by nucleotide, triphosphate co-enzymes, NTP such as ATP, and is feedback restricted. The environmental availabilities of these elements are not then directly connected to their distributions in cells which branch out in a vast array of managed compound formations throughout the cell. Furthermore the way in which the larger molecules are distributed in space depends on such factors as their solubility in water (and membranes) and strength of intermolecular interactions in complexes such that some aggregate in membranes, largely composed of H and C (fats), while others form assemblies, for example proteins, DNA and RNA which contain much N and O as well as H and C. To explain the way in which these individual elements are distributed in cells, biological organic chemistry [6,7] is clearly extremely difficult and it is then simpler to look at the distribution of some less common biological elements, biological inorganic chemistry [1–3], which deal with a small number of pathways of incorporation keeping in mind the basic needs of all the organic and most common elements to have several kinds of protein in their irreversible pathways, and for each of which it can be written:

$$\text{Selective uptake (on carrier)} \rightarrow \text{Carrier recognition} \rightarrow \text{Incorporation}$$

In higher organisms the uptake and distribution of the above elements is from larger compounds dissolved in aqueous solution, e.g. C/H/O from glucose, not from small units such as CO_2, N_2 and H_2O. In this case pumps or exchangers selectively take up compounds and metabolic control may rest at these stages, see below.

We can illustrate the problems by reference first to the distribution of sulfur which is central to metabolism of all organisms.

2.3. Sulfur distribution

Sulfur is most peculiarly distributed and dominantly situated in the cellular pathways of all organisms [6,7,13]. The major features of its distribution, Figs. 5 and 6, which affect biological chemistry and that must be noted are:

1. Most, perhaps not quite all, protein syntheses are initiated by a thioether amino acid, methionine, codon, so that methionine is initially the first amino acid of most proteins although subsequently it is often removed. Methionine synthesis is therefore a major step in sulfur distribution.
2. Many co-enzymes handling acyl and methyl transfer are based on thiols (cysteine) or thioethers (methionine) (see Fig. 5).

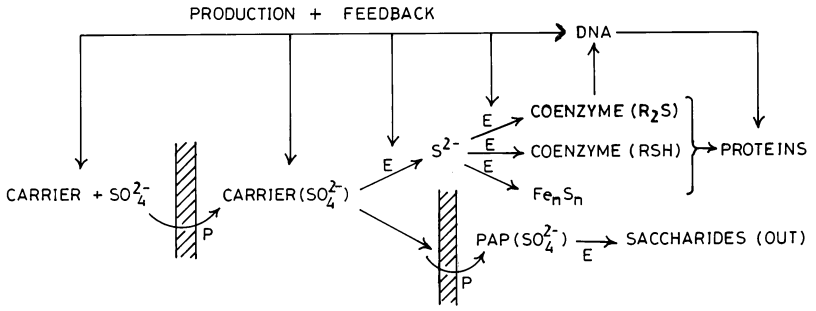


Fig. 5. A scheme for the uptake of sulfur and its distribution; E is an enzyme, P a pump, PAP is the sulfate coenzyme. RSH represents cysteine derivatives and R_2S methionine derivatives. Energy requirements are not shown. Membranes are shown as shaded zones, compare Fig. 3.

3. Much of the control of redox balance between reduced and oxidised organic material is managed via cysteine in glutathione (see Fig. 5).
4. Thiolate groups are at the active site of several enzymes.
5. Much cross-linking of proteins is due to –S–S– bridges.
6. N.B. (2) to (5) require sulfur to be distributed into cysteine as well as methionine.
7. The extremely important primitive and still existing energy pathways using redox electron transfer are based on Fe/S centres while H_2 uptake is based on Ni/Fe/S and Fe/S centres and N_2 uptake is based on Mo/Fe/S, V/Fe/S or Fe/Fe/S centres. Some sulfur is therefore retained as anionic sulfide, S^{2-}.

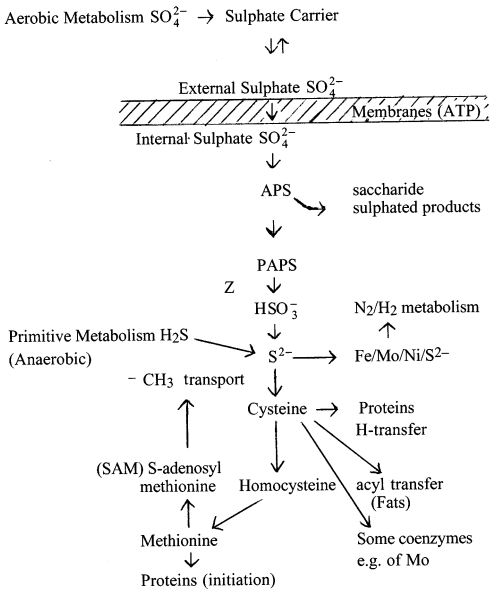


Fig. 6. The pathways of sulfur in many cells to give its very extensive distribution in many compounds using covalent intermediates. Feed-back connections are shown in Fig. 5.

Table 4
Major forms of sulfur

Sulfate	Incorporated into polysaccharides
Sulfide	Part of several metal co-factors, e.g. Fe_4S_4
Thiolate	Mainly in the amino acid, cysteine, and in co-enzyme Co-A
Thioether	Mainly in the amino acid, methionine and in co-enzyme S-adenosyl methionine (SAM)

8. Sulfate is also incorporated mainly in extracellular saccharides of higher organisms. Sulfate reduction also utilises compounds of iron and molybdenum involving sulfur itself.

Notice that sulfur has a major control over carbon, nitrogen, hydrogen and even its own incorporation, but this is achieved through distribution in only a few compounds. This dominance of sulfur was already present in the earliest organism due to its availability as sulfide and its reactivity. The cellular distribution of this element was made more complicated later in evolution by the oxidation of sulfide to sulfate. Even so sulfur is still present in a relatively small number of units throughout a cell, Table 4, which includes covalent attachment to carbon and ionic attachment to metal ions. We see this clearly when we examine sulfur uptake today which follows the assimilatory path shown in Fig. 6. Each arrow represents an irreversible enzymic process except for the first (extracellular) and third processes where transport to a membrane component is involved, but across a membrane a step is energised and irreversible.

The latter part of the pathway after S^{2-} formation is controlled by substrate feedbacks and regulated by gene expression, mainly based on MET, methionine, or SAM, S-adenosyl methionine, and cysteine (Cys) transcription receptors but it must not be thought that the feedback controls are simple. There are several related genes of each kind, well separated on the DNA. The possible complexity of the genes can be appreciated if we look at parallel genes for the initial incorporation of nitrogen which is a simple reaction pathway

$$N_2 \rightarrow NH_3 \rightarrow \text{glutamate}.$$

The gene structure for this process is shown in Fig. 7 [12]. Undoubtedly the initial handling of carbon is similarly complicated although the steps are again simple at first

$$CO_2 \rightarrow HCHO \rightarrow \text{sugars}.$$

This analysis of the distribution of sulfur has revealed many similarities with those for C, H, N and O but also one considerable difference. The distribution of these four major elements was based entirely on covalently incorporated units initially in co-enzymes. These elements entered covalent irreversible pathways immediately after diffusion through membranes as does H_2S. However appearance of sulfur as sulfate demanded carriers of this unit both within extracellular (higher organism) fluids and to allow it to cross membranes using pumps. Since the supply of sulfate is not automatically secure, like that of H_2O, CO_2 and N_2 from water and

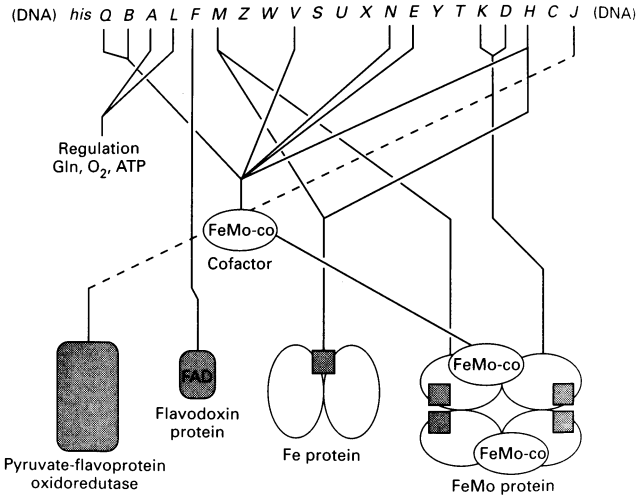


Fig. 7. *Nif* genes represented by italic capital letters required for nitrogen fixation as arranged in *K. pneumoniae* and their respective gene products. The regulation by product, glutamine (Gln), (dioxygen) and ATP (energy) is shown. This genetic structure for O_2 protection became necessary only after its advent. All genes except *A* and *L* generate RNA for the production of proteins directly or for the production of protein units for synthesis of FeMo-co.

air, respectively, methods were necessary to maintain it by capture devices. However the capture step due to reaction outside the cell could not depend on energy and irreversible covalent bond formation since it had to release the sulfate readily on demand from carriers at the pump and again on entering the cell. In fact this situation parallels that of all elements except H, C, N and O. Thus we find that sulfur, and all elements other than H, C, N and O, have initial uptake steps involving simple usually ionic equilibrium binding to (protein) carriers, exchangers and pumps. Equilibrium steps cannot be controlled but the activity of pumps can be managed by feedback. Hence pump, exchanger and carrier protein production became related to sulfate availability and had to be produced commensurately to ensure distribution in the organism. This is to say that since units such as sulfate (and many other elements) invariably supply the production of these proteins, the transport of sulfate into the cell had to be related to cellular demands. The necessity was then for the cell reproductive machinery, and for the code, DNA, to be informed of the level of sulfate in the cell and respond so as to regulate its uptake using carriers, exchangers and pumps. Hence we find an additional feature of

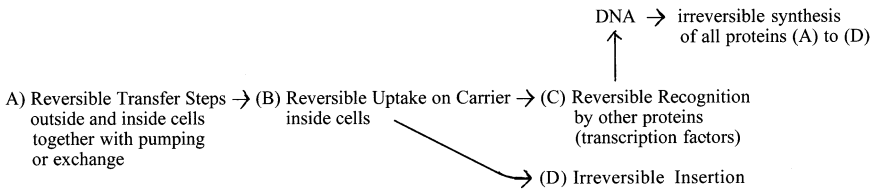


Fig. 8. Generalised pattern of uptake and incorporation in the cytoplasm.

distribution inside cells which occurs before irreversible combination in covalent structures and involves a distinct set of *equilibrium exchanges* between carriers, DNA, and the basic entity, here sulfate, Fig. 8 [13].

We see from this analysis that both enzymic pathways and extracellular and transmembrane transport steps govern the distributed incorporation of an element, here sulfur. The steps in the pathway can be switched on or off by different feedback systems so that sulfate uptake, sulfation, sulfide binding to metal ions, sulfur incorporation in methionine and cysteine, and co-enzyme production are all in harmony, both through feedback regulation of gene output of proteins and enzymes and due to the feedback controls over these enzymes and transport systems. Energy (ATP) is needed in pumping and later transfer steps of sulfate incorporation into sugars by phospho-adenosine phospho-sulfate (PAPS), the coenzyme carrying sulfate. The handling of sulfur is then further connected to carbon, nitrogen, and hydrogen incorporation and all of these are therefore homeostatically linked to sulfur systems. As a result of the management of the pathways of Fig. 6 sulfur is distributed in a fixed manner in a large variety of covalent compounds as well as in a few ionic complexes with metal ions. At the present time the network leading to the final distribution is again too complicated for analysis as was the case for H, N, C and O incorporation. The strong feature is that most, not all, sulfur compounds are covalent kinetic traps for the element, much as this is true for C, N, O, and H. A very similar regime applies to selenium. One further complication is that some of the reactions do not occur in the cytoplasm, for example sulfate incorporation in saccharides takes place in the Golgi, implying further equilibrium ion exchange steps at the pumps before the final covalent binding.

Before going further observe that of the six elements mentioned so far, four, C, N, O, and H, are present in DNA (RNA), and all six are coded in covalent amino acids and that no other elements are so coded. Thus their uptake and incorporation using enzymes, pumps and carrier proteins is largely *constitutive* but it is product regulated and can be enhanced. Due to the resultant complications of their distribution we shall look for understanding of element distribution in the yet again simpler incorporation of an element concerned with the handling of sulfur, molybdenum. This is the next convenient element for the analysis of distribution since its pattern is similar to but simpler than that of C, H, N, O, S or Se. Note that nitrogen incorporation requires molybdenum and we shall see that molybdenum uptake/rejection is based on sulfur chemistry, see below, while sulfate metabolism depends on molybdenum. Already we see again and again the interwoven pattern of element dependencies. The implication is that eventually we must deal with complex system analysis, not with simple sums of separate steps of individual elements, but reductive analysis of the individual element reactions is at present our best and only way in which to proceed.

2.4. Molybdenum incorporation [14,15]

As stated, in an effort to simplify the discussion of distribution, starting from the major elements it is convenient to look next at molybdenum, and the proteins which handle it, since it is only found in two final coenzyme products. The related

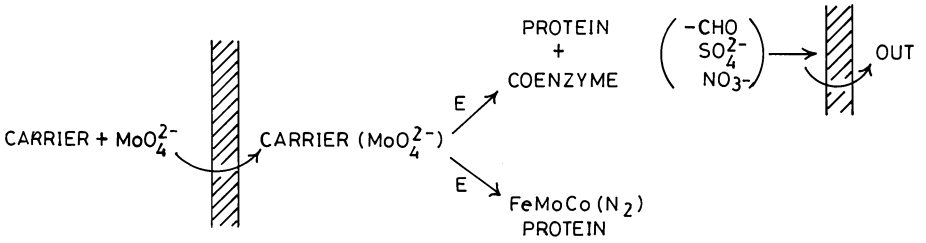


Fig. 9. An outline of molybdenum uptake and distribution which is parallel to Fig. 5. Energy requirements are not shown. In brackets are the substrates of the enzymes which are often exported.

genes and proteins are also relatively simple and are referred to by the label Mod. A third reason for discussing molybdenum next is that there is some similarity to the genes and gene protein products for the initial steps in handling of molybdate with those for sulfate, Fig. 9. Of course the compounds have formulae which are similar, SO_4^{2-} and MoO_4^{2-}, and so are their carrier proteins. The molybdenum carrier, Mod A, found in the extracellular fluids of higher organisms has in fact a fold much like that of the SO_4^{2-} carrier and transferrin, the iron carrier [14]. Somewhat curiously and differently, bacteria synthesise special small molecules for Mo uptake which are similar to iron siderophores [15]. Note the similarity to Fe^{3+} uptake molecules in both prokaryotes and advanced eukaryotes and also the similar chemistry involved [16]. The carrier gives MoO_4^{2-} to a membrane ATP transporter, Mod B, a pump which is open to feedback control, Fig. 10. Before molybdenum is modified in the cytoplasm to give the co-enzymes associated with aldehyde, nitrate and sulfate metabolism and with nitrogen fixation, molybdate is carried by a further transporter protein, Mod C, in the cell. These transporters may well act also as stores for MoO_4^{2-}. Now all these proteins must have their production regulated at the DNA level and there is a known Mod E protein which binds MoO_4^{2-} and

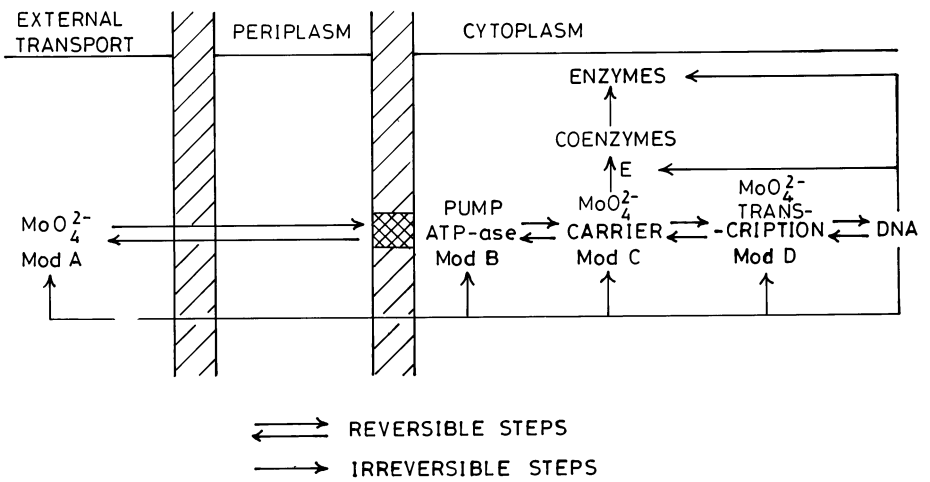


Fig. 10. A more detailed analysis of the uptake of molybdenum into *E. coli*. The irreversible steps have to be controlled at enzymes or regulated at DNA by feedback which is not shown, see Fig. 3.

DNA, and informs the genetic machinery of the adequate (Mo bound to Mod E) or inadequate (Mo not bound to Mod E) supply of the element. At this level the molybdate anion must be in a *balanced ionic equilibrium* between binding to the signalling system and to the carriers while it acts as a feedback control over incorporation and uptake systems to maintain and secure management of its distribution (see Fig. 10). All the Mod proteins have to be able to distinguish the molybdate anion from SO_4^{2-} and HPO_4^{2-} (see Section 3.5). In comparison with the handling of sulfate, the number of ionic equilibrium steps is now considerably increased over those of irreversible incorporation which follows them.

The next proteins to bind molybdenum are the enzymes of nitrate/sulfate reduction, see step Z of sulfate metabolism, Fig. 6, and of Mo reduction to give FeMoco, see steps QBNE of nitrogen fixation, of the nitrogen fixation genes, Figs. 7 and 9. These steps involve (energised?) irreversible (covalent) binding into two types of co-enzyme. Thus the handling of molybdenum in a cell involves a series of reversible transfer or transport, anion-binding steps ending in irreversible insertion into enzymes [17]. In some organisms the insertion of one of the co-enzymes occurs outside the cytoplasm in the periplasmic space. Many of these steps require the application of energy (ATP). The element distribution is finally like that of H, N, C, O, S and Se in that molybdenum is to a large extent irreversibly covalently trapped in proteins or organic co-enzymes. The distribution of the element is then governed in ways simpler but similar to that for sulfur and indeed to that for selenium. Thus despite the fact that we have examined a system of very much greater simplicity than that for say carbon, (only two molybdenum compounds are formed and in most organisms only one) we have not yet been able to appreciate fully the pathways and controls to even such simple products. The problems are the initial control mechanisms for uptake using equilibrated exchange, the switches in oxidation state, covalent incorporation in late irreversible steps, co-incidental production of organic and inorganic centres for binding molybdenum and production of all the proteins concerned as feedback regulated gene products. Note again that the amounts of the organic pterin part of the Mo coenzyme, of FeMoco, and of their corresponding proteins are all produced so as to leave very little excess of any portion of them not bound to molybdenum. As we stated previously, the balanced production of partners is a very general feature of cellular feedback controlled homeostasis, but note also the transfer modes across membranes (Table 5).

Table 5
Element transfer modes across membranes

A. Energised transfer	
Coupling to ATP-hydrolysis	ABC-transporters
Coupling to ion gradients	Na^+ or H^+ exchangers
Coupling to electrical potential	H^+ flow
B. Simple transfer	
Channel openings	Gated (electric potential or chemical) channels

2.5. Incorporation of phosphate: an aside [4–7]

In order to describe the handling of all the non-metals which are covalently bound before discussing the metals we insert an aside to phosphorus.

Phosphorous incorporation is in one respect even less complicated than that of all the major elements H, C, N, O and of the elements mentioned above, S, Se and Mo, since phosphorus is only described in one oxidation state, phosphate. Phosphate is processed in much the same manner as the sulfate and molybdate in the steps leading to entry into the cytoplasm. However after entry it is immediately incorporated into a huge variety of organic esters and anhydrides although it is never reduced. The distribution of phosphate is in fact largely in irreversible covalent compounds, which are connected to transcription factors and to pumps across membranes, and hardly any steps, many of which require energy (ATP), involve reversible ionic equilibration except in the initial stages of uptake and transfer. Note that as with the other elements there is not only internal control and regulation of the different steps involving phosphorus but there must also be cross-linkage to the control and regulation of the distribution of the other elements, and there is a demand for energy. We have seen vestiges of this feature particularly in the diagram of the nitrogen fixation genes, Fig. 7. In fact phosphate metabolism is seen to be connected to that of all the elements we have described earlier through nucleotide triphosphates, NTP. It is observed that the levels of nucleotide phosphates, e.g. ATP, and phosphate are fixed for a given metabolic state of a cell. Since ATP carries the energy for the cell and it acts as a feedback control on a very large number of reaction pathways through the reversible binding of ATP to pumps, kinases, signalling proteins and so on, it had to be of a virtually fixed affinity $K \simeq 10^3$ to all proteins in the pathway. This helps to ensure the coordinated activity of primary metabolism in all cells — homeostasis. We shall find that distribution of an element amongst different centres, when close to equilibrium as in the case of ATP, requires very similar binding constants at all its binding sites. Moreover the functioning of ATP depends on the binding to a metal ion, Mg^{2+}, at equilibrium, which implies that $[Mg^{2+}]$ had to be at a fixed concentration too [20].

Once again the distribution of phosphorus is extremely complicated and under the regulation of a large number of feedback gene products and enzymatic control processes.

2.6. Competition between ionic species for binding centres [2]

During the description of C/H/N/O incorporation and distribution we only dealt with highly specific irreversible kinetic steps due to the recognisably different shapes and sizes of covalent molecules and anions, including the initial carrier co-enzymes. However when we came to examine the distribution of the other non-metals, S, Se, Mo and P, several initial steps involved the reversible ionic binding of SO_4^{2-}, SeO_4^{2-}, MoO_4^{2-} and HPO_4^{2-} all of which have similar shapes, and hence can compete with one another. In these steps we must analyse the ability of the carriers to discriminate between the different anions. It would appear that equilibrium

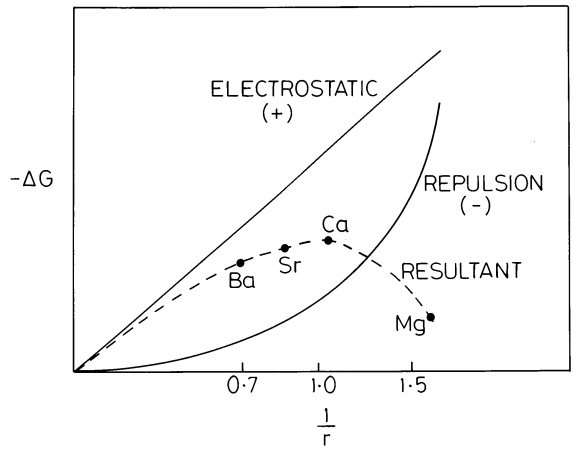


Fig. 11. Ionic binding to ligands is based on long-range electrostatic attraction and short-range repulsion. Depending on the ligands the resultant binding free energy may be at a maximum at any given size due to the repulsive term or to constraints on the size of the cavity the ligands can generate.

selection is largely based on size differences (see Section 3.5) where there is little difference in shape. Next we shall turn our attention to the parallel difficulties in the separate distribution of cations where the stress is now mainly on equilibrium exchange. These ions are usually handled at first by transport at *ionic* equilibrium, with little or no covalent binding, and for many elements this is also true of their final state of incorporation. We shall see if we can uncover truly simple systems of element distribution in cells based upon thermodynamic equilibria, themselves to be appreciated through the analysis of simple model systems, Section 3. Note that equilibrium systems are not open to control so that since we know that the levels of all elements are controlled we need to find managed steps. In fact only where energy (ATP) is applied, e.g. at pumps, are the steps irreversible and open to control. Hence for balanced production of their proteins, regulated by the free ion level, the pumps must have feedback control by this same level [3]. For the simplest cations equilibrium binding selection is again based on size (Fig. 11) and charge [1,2].

2.7. Incorporation of simple inorganic ions

Firstly, most of the incorporated ions are metal ions (Na^+, K^+, Mg^{2+}, Ca^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^+ and Zn^{2+}) while there are but a few anions (Cl^-, Br^- and I^- and those described above). They are handled by simple ionic transport processes outside cells and across membranes (including exchange or pumping). The transfer is of bare ions or reversible chelate complexes and not usually by enzymic covalent reactions. A particular extreme situation very different from that of common organic elements (H, C, N, O in compounds) or from that of molybdate, selenate, sulfate or phosphate, arises for elements which are present as very simple ions such as Na^+, K^+ or Cl^- and where the distribution of them in any one phase

is certainly directly related to the concentration of their free ions since the free ions there are clearly in equilibrium with any bound forms. The uneven concentrations of the elements inside and outside the cells or cell compartments (including vesicles) is then directly related to the protein machinery, pumps, exchangers and channels, which control them by putting energy into their ionic distributions [18]. It has to be remembered that these elements as ions are in high concentration and can diffuse freely and since few of them are bound, they do not need carriers for their distribution. Thus we need to know only what causes their channel and pump proteins to be expressed, and what controls the energy put into these devices. The extreme case is of the distribution of an element always in the same (ionic) state (see Fig. 12). There are four limitations on distribution: (i) energy (ATP) supply; (ii) feedback blocking of the inward pumps; (iii) control of exit mechanisms; and (iv) regulation of pump and channel production. In Fig. 12 we have included two irreversible steps to ME_1 and MP_6 apart from protein production.

Of course even in the most primitive cells much of the uptake or rejections of these elements are essential activities and hence the pump or exchanger or channel protein expressions have to be constitutive, not induced. The uptake/rejection modes are essential to the life of all cells. The *primitive* protein systems for these flows of ions were either exchangers for metal ions across the outer membrane or pumps driven by energy generation from proton gradients or ATP from metabolism. Since much ATP production is itself dependent on proton gradients one likely source of initial energy was the proton gradient itself [19], where the proton is an example of a simple cation. The overall implication is that the ATP and proton energy levels and exchange activities will come into controlled feedback balance with the simple ion gradients of Na^+, K^+, and Cl^- to maintain a viable cell in osmotic and electrolytic balance with the environment. The feedback control by the ions is exerted at the pumps and channels. So long as the pumps and channels are constitutively produced there is no problem in achieving a fixed

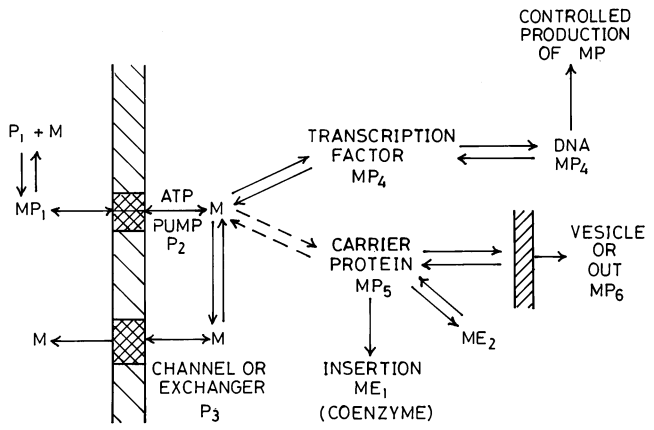


Fig. 12. Uptake steps of ions. Irreversible steps involve all pumps (internal transfer in membranes) and protein productions at RNA. All other steps, except to ME_1 and MP_6, are drawn as reversible, contrast Figs. 4–10. The figure applies particularly to K, Na and Mg.

balance point by feedback on these proteins. The balance is dynamic in that ions constantly flow in and out of cells. The distribution of the elements is then understandable from the energy applied to create it using feedback at a fixed equilibrated value in each separate compartment. In primitive organisms direct binding of these three elements, Na^+, K^+ and Cl^-, to proteins which bind DNA and act as reversible transcription factors is not apparent. Potassium is also involved with certain kinases and in DNA telomere structure where the bound condition equilibrates with the free. Involvement of sodium or chloride in metabolic processes is less well documented.

2.8. Selective incorporation of ions

As in the cases of SO_4^{2-}, SeO_4^{2-}, MoO_4^{2-} and HPO_4^{2-} there is a requirement for selective interaction with proteins so that Na^+ is distinguished from K^+ and Cl^- is not confused with simple anions such as OH^-. Generally these four ions, Na^+, K^+, OH^- and Cl^- are sufficiently different in their charge and size, and there are no other common monovalent ions present, to allow easy recognition (see Fig. 11). Thus in any one compartment of a cell or a higher organism there is found to be a fixed background distribution of these four free ions plus some localised spots of weak binding, e.g. K^+ in the telomere of DNA, and close to membranes where the negative charge density is high. All sites are in equilibrium.

2.9. Pumps

This is a good point in the discussion to insert a note concerning pumps. A pump delivers an element (or compound) across a cell membrane up to a certain external/internal gradient. The binding to either face of the pump is at equilibrium with the concentration of the element in the solution adjacent to that side, however one side of the central action point of the pump has a different affinity from the other due to energy applied irreversibly at this point. In the case of H^+ pumps, ATP-synthases, the energy applied is close to being reversible. The application of energy to the pump, then moves ions (molecules) against the gradient irreversibly and is often controlled by a further separate binding site for the element such that at a given concentration of the ion inside the cell the pumping action is stopped by allosteric feedback.

A much more complicated distribution has to be made in higher organisms especially multicellular plants and animals. Here the local internal environment, the extracellular fluids, have to be maintained as well as the intracellular media with a variety of separated compartments (vesicles). The roots of plants and the digestive systems and the kidneys of higher animals provide the control over the extracellular fluids. The overall ion levels are now managed in part by messengers (see Section 3.3) — sterols from glands for example — which act on transcriptional regulation of pump expression in particular organs. Thus there is a degree of local enhancement as well as of constitutive expression. Inside the cells separate pumps,

exchangers and channels allow distribution in the cytoplasm and the vesicles to be managed.

2.10. Energy of membrane uptake and exit of elements

In the previous sections we have often stated that passage through membranes is a required feature of the uptake or rejection of elements from or to the external environment and to vesicles (see Figs. 3 and 12). These transfers require that there is an initial favourable gradient or that energy is used in pumping against the gradient. Table 5 lists the transfer modes.

The nature of biological cells requires that all the processes in Table 5 should be selective following the principles outlined in previous sections and that all of them should be under feedback control.

2.11. Exchangers and channels

Channels are just selective filters allowing almost specific flow through membranes. They can be gated so that the flow is activated by an electric potential or by chemical binding at a receptor point on the channel protein. Exchangers allow inner passage of one ion balanced by outward passage of another. They can be electrically neutral or not when they do or do not create a potential gradient and are then energy coupled. Typical neutral exchangers are $2Na^+/Ca^{2+}$ and $2H^+/Ca^{2+}$. They and indeed some other channels, much like pumps, are under feedback control, gating, from the very ions which pass through them where one ion, e.g. H^+, is produced inside the cell to generate an initial energised gradient. H^+, pH, equilibrates rapidly in all compartments except organelles.

2.12. Magnesium and calcium distributions [20,21]

The next group of elements of concern to us are those that also exchange rapidly but bind somewhat. They are Ca^{2+} and Mg^{2+}. Again they are not involved in covalent binding and they are not trace elements. The case of Mg^{2+} is the simpler in many multicellular organisms since its concentration, close to 10^{-3} M, allows rapid free diffusion to equilibrate this ion in each separate compartment, including inside the cytoplasm, outside cells, and in organelles although not into vesicles. It is unevenly distributed in a few vesicles. Notice that Mg^{2+} is not usually pumped into vesicles, see Fe^{2+} and contrast Ca^{2+} or Cu^+ and Ni^{2+}. Some control is clearly required only over Mg^{2+} inwardly and outwardly directed pumps and channels which are constitutively expressed. Regulation plus control implies both the synthesis of these proteins and feedback limitations on their activities. We can therefore deal with distribution by reference to thermodynamic binding of Mg^{2+} ions to their binding sites in each compartment (see Fig. 24) together with the energy applied to the crossing of each membrane. As already mentioned, a quite

striking exception to this generalisation is the appearance of magnesium in chlorophyll where it is kinetically trapped (see Section 8).

The situation for bacteria or any other cells in fresh water of low Mg^{2+} content is very different [20]. They must have an energised uptake of Mg^{2+} to maintain the universal cellular concentration of 10^{-3} M. In fact it is known that such cells use Mg^{2+} pumps so that the concentration of 10^{-3} M is maintained but not exceeded. Thus as for the simple ions Na^{+}, K^{+} and Cl^{-} distribution is maintained constant by supplying energy to pumps and channels and control over them by feedback. Whatever the devices they are again to a large extent constitutive and not induced.

Mg^{2+} now binds to many transcription factors as well as to RNA and DNA directly through phosphate groups, which undoubtedly affects the cellular activity. It also binds to enzymes which modify transcription factors such as kinases and phosphatases. In fact much ATP is bound as MgATP. Since Mg^{2+} equilibrates in the cell the distribution is grossly simplified by having all its binding constants in the range 10^{3}–10^{4}. A further problem which arises is the management of this distribution in the face of competition from many other, more strongly binding, cations including Ca^{2+}, Mn^{2+}, Fe^{2+}, and Zn^{2+}. We shall need very careful analysis of effective binding constants to reach some appreciation of how all these ions are separately distributed (see below and Sections 3–8).

The case of calcium is different again [2,21]. The cation exchanges rapidly but its low concentration 10^{-7} M in the cytoplasm of very many cells makes it difficult to maintain its distribution, much like magnesium, virtually all its interactions in a given compartment can be described by thermodynamic equilibria. There are also proteins for the control and distribution of free calcium in eukaryote cells in addition to the pumps, exchangers and channels for inward and outward transport across outer membranes so that calcium is distributed in many vesicles at about the same concentration of 10^{-3} M. In higher organisms the circulating levels of calcium in the extracellular fluids of 10^{-3} M is also very well controlled. It appears that it is linked to the level of phosphate and to that of pH so that materials such as shell and bone are formed in a well-defined manner.

As is well-known calcium is the major second messenger of many cells so that its fast entry into the cell on receipt of a primary message generates a novel and temporary new state of the cell opposite a new internal calcium concentration $> 10^{-6}$ M. The change in state has to be coordinated in the many altered cellular activities. The calcium ion then acts (almost simultaneously) on a set of proteins, maybe as many as 40 in advanced cells and these include mechanical switches (troponins), membrane adjusters (annexins), transcription factors (calcineurin) and metabolic switches especially acting on kinases (calmodulins and S-100) [2]. Calcium levels therefore act in cooperative and coordinated action, Fig. 13 [2,3,21]. To do this calcium binding constants to every centre in the cytoplasm must be very similar, around $K = 10^{6}$–10^{7}. This too must be slightly less than the binding constant of the inner part of the calcium pump. Thus calcium acts as a coordinating agent for an excited state of the cell of short duration, and a homeostatic agent for longer lived changes of state.

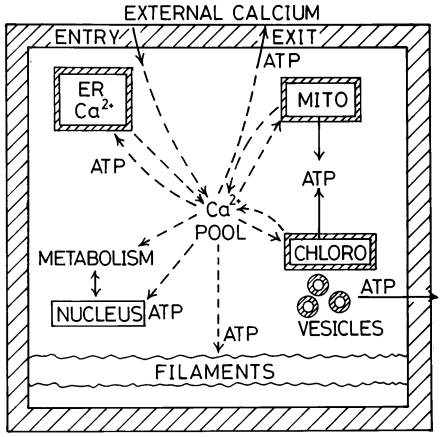


Fig. 13. The dynamic distribution of Ca^{2+} ions in eukaryote cells means that it acts as an electrolytic message carrier integrating many features much as an electron operates in an integrated electronic circuit.

Now the time response to the calcium message can be modulated by the rate of diffusion of calcium in the cell-itself due to calcium carriers — calbindins. Here the binding constant must be the same but the exchange rates are slower by up to 10^3 [21].

As stated inside and outside the vesicles of a cell of higher organisms which responds to calcium there are 10^{-3} M free calcium. A variety of proteins there bind calcium with a binding constant of 10^3–10^4, Fig. 14 [22]. The homeostasis of these spaces is therefore very differently set from that of the cytoplasm, the calcium ion now acting as a coordinating agent amongst active proteins ejected from cells [21].

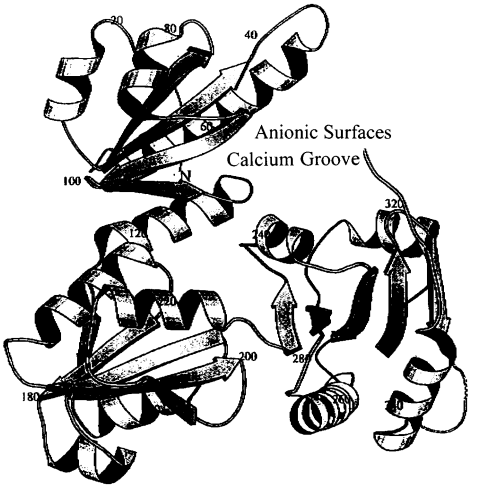


Fig. 14. The structure of calreticulum [22]. It can bind 30-40 Ca^{2+} ions with an almost fixed binding constant of 10^4.

To maintain this system requires access to calcium from the environment and a maintained level of proteins, which bind calcium, especially in multicellular organisms. Several of these proteins are constitutive but others need to be enhanced during growth. In fact the level of calcium circulating and in (close to) equilibrium with minerals, e.g. bone is directly linked to the hormonal levels of certain steroids. This is also true of K^+ and Na^+ levels. These steroids, secreted by glands, act on epithelial and bone cells so as to control levels of calbindins, the carrier of Ca^{2+} in epithelial cells, and of osteocalcin, which controls bone dissolution. Thus while the whole system acts as a series of linked equilibria in vesicles, in the cytoplasm and in the extracellular fluids the linking can be modulated by Ca^{2+} feedback at pumps or by regulating the synthesis of the Ca^{2+} uptake transport systems and finally by the manipulation of minerals held close to equilibria with free ions. The well-being, including growth of the whole organism, is then linked to the setting of the calcium circuitry, Fig. 13, and well known diseases are due to calcium defects, e.g. rickets [21].

Now all of this calcium chemistry, at separate equilibria in each compartment, has to avoid other metal ion competition. This is managed by the type of chelating agent the cell generates. The Ca^{2+} protein binding sites are all O-donors and their coordination number is greater than six (see Fig. 19). The bulky nature of their binding cavities makes them poor for binding Mg^{2+} and other small cations Fig. 15 but the only reason they do not bind other trace elements such as Zn^{2+} lies in the affinity of these elements for N- or S-donors, which removes them (see Section

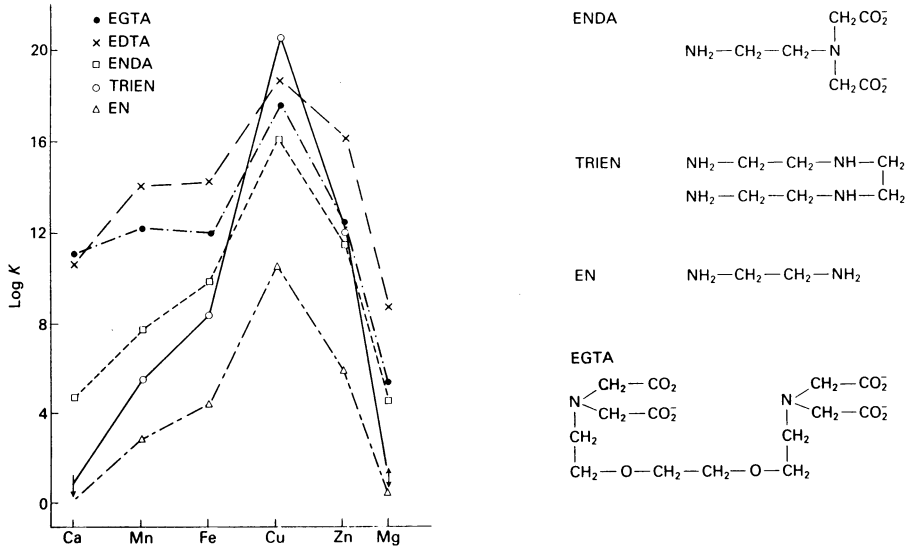


Fig. 15. The stability constants, log K, of some complexes of the elements in their divalent M(II) states compared with Ca(II) [23]. The symbols refer to ethylenediamine, EN; triethylenetetramine, TRIEN; ethylenediamine diacetate, ENDA; ethylenediamine tetracetate, EDTA; and ethyleneglycoldiamine tetracetate, EGTA. Note how the increase in O donor groups stabilises Mn(II) and Ca(II) relative to other elements.

4) and so reduces their effective stability constants with the calcium ligands. To understand the necessary effective equilibrium constants in a cell we shall look at model systems of binding for a range of cations in Section 3.

2.13. Heavy trace element distributions

The elements we introduce next are zinc and copper since neither of these elements is found in large irreversible co-enzyme chelates — contrast Fe (in heme), Co (in vitamin B_{12}) and Ni(in F-430). The situation is now one degree more complicated since some of the proteins which bind these metals are in effect in irreversible association, returning us to the type of description we used in discussing sulfur and molybdenum distribution. In other words the metallo-proteins are the entities to be considered not the free metal ions since these metallo-proteins have concentrations much greater, by several orders of magnitude, than the free metal ions. Again these two metals and their proteins are unevenly distributed in vesicles as well as in different cell types.

The cases of Fe, Co and Ni are even more complicated due to their incorporation in non-exchanging sites of the above co-enzymes as well as more generally. Due to the additional complications of competition between their free ions and Mn^{2+} and Mg^{2+}, for many proteins we shall examine all their ionic equilibrium distributions together in Sections 3 and 4 before we deal with many enzymes in Sections 5–7.

Note that just like sulfur the destination of metals such as Cu, Zn, Ni, Co and Fe can be to both exchanging and *non-exchanging* sites which makes the problem of their description more difficult than that for Na^{+}, K^{+}, Mg^{2+} and Ca^{2+}. In fact the problem of distribution becomes more and more like that seen for sulfur (H, C, N, O and Se) as we pass along the series of increasingly strong binding to proteins.

Ionic binding very weak	Ionic binding weak	Binding moderate	Increasingly strong binding	Covalent
Na^{+}, K^{+}, Cl^{-}	Mg^{2+}, Ca^{2+}	Mn^{2+}, Fe^{2+}	Co, Ni, Cu(Zn)	H, C, N, O, S, Se, Mo
	Exception Chlorophyll	Exception Heme	Exceptions B_{12} F-430 and some proteins	

A major problem now arises even in situations of fast exchange in that the whole series of ions Mg^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Co^{2+}, Ni^{2+}, Ca^{2+} and Zn^{2+} are all relatively similar in binding capability to certain binding groups so that biological systems have extreme difficulty in achieving selectivity. One solution to the problem is a variety of reversible equilibrium steps of great complexity. We shall attempt in the next sections to generalise this approach to managed distribution in order to understand why different elements are found in association with different organic groups especially proteins or in different compartments. Of course in the final analysis all these distributions are linked to functional value. We shall relate the analysis to the extent to which protein production systems are constitutive or

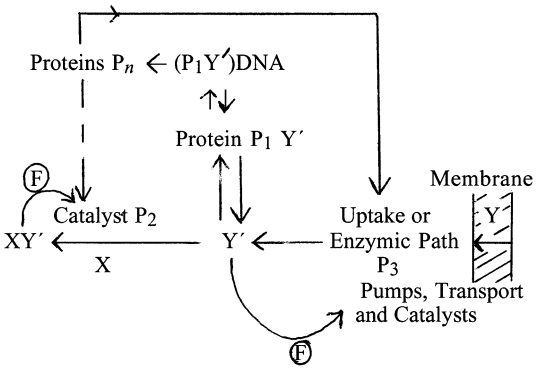


Fig. 16. The nature of pathways, $Y \rightarrow Y' \rightarrow XY'$ (full-lines) and their connection to protein production (dashed-lines) showing also feedback (F) relationships. Note equilibria are shown ⇄. This is a simplified version of Fig. 3. P is a protein.

induced, since the management of distribution relies on limited access of different metal ions to limited production of selected proteins. Where the equilibrium treatment fails we shall return to kinetically controlled exchange.

3. Basic steps of distribution and equilibrium reactions: model systems [1]

3.1. Introduction

We started this discussion of the distribution of the elements in cells with an analysis of the distribution of H, C, N, O, and in more detail, sulfur. The distribution is for the most part controlled by the simple principle that once an element is incorporated into a specific *covalent* organic compound it can be handled *irreversibly* and very *selectively* since the compound has a recognisable molecular shape. The further application of energy then allows further irreversible incorporation into more and more complicated molecules and into different compartments. All compounds are also degraded by other enzymes or simply because they are energised. Feed-back control, again utilising recognition of molecular shape, then generates a balance between synthesis and degradation of small molecules and proteins by acting upon enzymes. There are further feedback controls, now regulating production of those proteins, that act on the pathways of incorporation of the elements via transcription factors which affect DNA. To be effective each feedback must be in fast exchange relative to the process it manages, that is close to equilibrium. Such feedback equilibria only apply to some of the later binding reactions of free metabolic products, that is after incorporation of H, C, N, and O into molecules of molecular weight over say 100 Da. Examples are the binding of sugars or glutamine to their transcription factors. Similar equilibria also apply to later steps in sulfur metabolism after it has been covalently incorporated, e.g. into methionine, so that each pathway is regulated both through feedback by its own

products on its own enzymes and pumps and upon regulation of the production of these enzymes and pumps (see Fig. 16). So it was in all metabolic pathways of H, C, N, O, and S (Se) chemistry with but one or two important exceptions. The initial steps of sulfur uptake as sulfate required additional pumps and carriers so that sulfate binding is in an *ionic* equilibrium, both to carriers outside the cell before incorporation, and to the outside surface of pumps, i.e. there are no initial irreversible covalent steps. Inside the cell itself after pumping, which is irreversible, the handling processes of ionic sulfate is also in equilibrium contact with the inner surface of the pump and with gene controlled production of proteins. This is to say that an *equilibrium* must exist between the sulfate bound (to DNA control proteins — transcription factors) and free sulfate. The distributions under initial irreversible metabolic reactions are difficult to analyse but those steps in which ionic equilibria are established, as in the case of sulfate and which as we have stated are quite common in metal ion chemistry as well as for some other non-metals or metal in anions, e.g. molybdate, are easier to appreciate. The following sections will therefore start the analysis of the distribution of elements in the cytoplasm of cells from the opposite extreme of behaviour, ionic equilibration, relative to the procedure used in Section 2 where we started from consideration of covalent incorporation. Here we meet the problem that equilibrium involving one ion or molecule is open to severe competition from similar ionic entities and it is the resolution of this competition which we analyse first using model complex ion reactions to appreciate selective metal ion/ligand interactions at equilibrium [1,2]. We conclude with a brief reference to anion uptake.

3.2. General principles: equilibria in model complexes

In order to understand the distribution of a large number of metal ions, approximately ten, amongst a wide variety of metal binding ligands where ligand and metal ion concentrations are somewhat equally limited as in cells so that under conditions of strong binding there is present very little of either free ligand, X, or free metal, M, although MX may be present in considerable amounts in the system and each MX is selectively formed, we can analyse the stability constants for complexes of simple model ligands which bind selectively to the metal ions concerned under conditions of low availability of free ligand. We first take the analysis of bisbidentate chelate complexes in which four donor atoms bind divalent metal ions but which allow any stereochemistry, Fig. 17. There are several factors of interest:

1. Multi-carboxylate anions, represented here by two oxalate units, Fig. 17, bind to all divalent metal ions with an affinity constant greater than 10^4 and less than 10^8 ($\log \beta_2$). They are charged O-donors. The same situation applies to ATP 1:1 complexes, Fig. 18.
2. Amongst neutral donor selectivity, the slope of the plots in Fig. 17 increase from O-donors to N- or S-donors
3. The slope due to four S donors (bis-2,3-dimercapto, 1-propanol), 4N donors (bis-ethylenediamine) or 2N, 2S donors (bis 2-mercaptoethylamine) is greater than that due to 2O, 2S donors (bis-β-mercaptoacetate), or 2O, 2N donors (bis-amino acid complexes), or for two 4O-donors (oxalate and ATP).

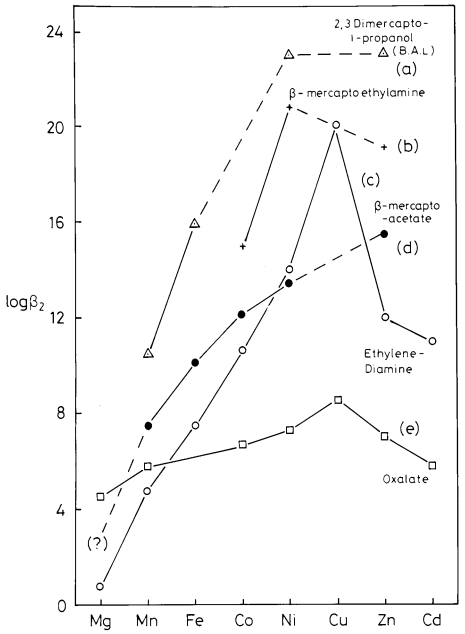


Fig. 17. A comparison of the stability constants of ligands using S→, N→, and O→ donors. Data from [23].

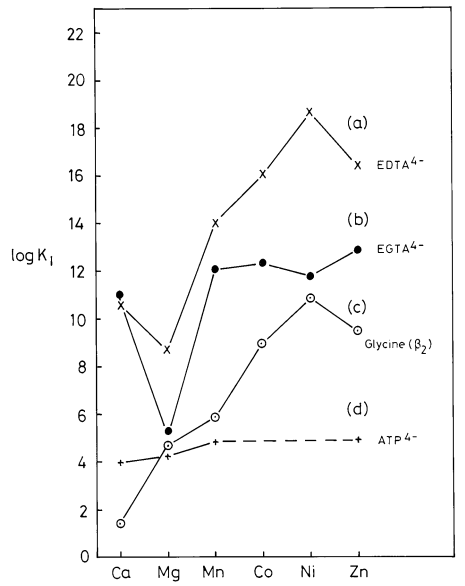


Fig. 18. A comparison of the stability constants of ligands using O→ donors or O→ and N→ donors where the ligands have various numbers of donors and packing constraints (see Fig. 15). Data from [23].

Table 6
Stability constants of $M(X)_2$ where X is ammonia or imidazole [23]

	log β_2									
	Mg^{2+}	Mn^{2+}	Fe^{2+}	Co^{2+}	Ni^{2+}	Cu^{2+}	Cu^{+}	Zn^{2+}	Cd^{2+}	Ag^{+}
NH_3	0.3	1.5	–	3.7	5.1	7.8	10.6	4.9	4.9	7.2
Imidazole	(0.3)	3.0	2.3	4.4	5.5	7.7	10.7	4.9	5.0	6.9

4. Certain donors have a high affinity for protons. As a consequence the effective *binding constant* of each RSH group is some 10^3 lower at $pK = 7$ than is shown in Fig. 17. Binding to RNH_2 ($pK_a \sim 10$) groups is almost equally affected by pH but in proteins RNH_2 is not a usual N-donor and is replaced by imidazole which has a pK_a close to 7.0. Now imidazole binds to metal ions with a strength very similar to that of ammonia Table 6, and thus the line in Fig. 17 for two ethylenediamines is probably quite closely representative of the binding constant of 4N-donation by four nitrogen atoms of imidazole. Due to H^+ competition Mn^{2+} and Mg^{2+} are not expected to bind to S-donors at pH 7.
5. To a first approximation RS^-, imidazole and RNH_2 are equally discriminating, that is they have approximately equal slopes, e.g. compare $NH_2{\cdot}CH_2CO_2^-$ and $^-S{\cdot}CH_2{\cdot}CO_2^-$, but in absolute log K terms the second is the stronger ligand. They probably bind with somewhat equal strength at pH 7 to the strongly binding cations, Zn^{2+} and Cu^{2+}.
6. Although (3) and (5) are true to a first approximation there are individual preferences amongst metal ions. Thus (ethylenediamine)$_2$ binds Ni^{2+} 100 times more strongly than Zn^{2+}, while Zn^{2+} binds (β-mercaptoacetate)$_2$ 30 times more strongly than Ni^{2+}. Thus a separation and hence speciation of nickel from equal amounts of zinc is clearly possible at equilibrium provided the supply of each of the two ligands is restricted almost exactly to the amount of each metal ion present, a condition we expect to hold in a cell for strong binding (see Section 1.5) with feedback limitations on free [M] and protein concentrations.
7. The change of valence of copper to Cu^+ greatly increases the affinity of copper for RS^- and somewhat increases affinity for N-donors while the change of valence of iron to Fe^{3+} greatly increases iron binding to S^{2-}, O^{2-}, OH^-, and RO^-, e.g. phenolate or carbamate.

Generally from the above data we consider the following is likely to be true for divalent ion binding to donor centres at equilibrium where the amount of a given type of donor is not greatly in excess of the individual metal concentrations to which it binds best.

	Mg^{2+}	Mn^{2+}	Fe^{2+}	Co^{2+}	Ni^{2+}	Cu^{2+}	Zn^{2+}
Donors	**O**	**O**(N)	**O,N,**(S)	**O,N,**(S)	(O)**N**(S)	**N,S**	(O),**N,S**
			↓			↓	
			Fe^{3+}			Cu^{+}	
			O,N,(S)			**N,S**	

As stated copper and iron can also be removed from this competition by valence changes, which give rise to special binding peculiarities.This selection of ligands and

Table 7
Radii and ionisation potentials (*IP*)[a] of some elements [1]

	Ca	Mg	Mn	Fe	Co	Ni	Cu	Zn
Radius M(II) (Å)	0.99	0.65	0.78	0.72	0.69	0.65	0.65	0.65
IP M → M(II) (eV)	11.87	15.04	15.64	16.18	17.06	18.17	20.29	17.96

[a] The ionisation potential, $M \rightarrow M^{2+}$, is the same as the electron affinity of the reverse reaction $M^{2+} \rightarrow M$ and is a good measure of the acceptor power of the ion, M^{2+}. Similar tables can be given for M(III), etc. The order of 'softness' is related to *IP*.

binding strengths reflects the physical parameters of the ions given in Table 7 and is related to the Irving–Williams series $Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$. To a good approximation this distribution is observed in living organisms (see Section 4).

We turn next to the additional selection imposed on this order by stereochemical factors due to constraints on ligand geometry. The particular demands of individual cations have been listed in Table 8. Before doing so notice that especially Fe^{2+}, Co^{2+}, Ni^{2+} and Zn^{2+} have very similar demand preferences for particular ligand donors. Hence, to gain additional selectivity of one of these metal ions over the others, structural factors as well as electron affinity have to be used (see Table 8). The most common are the size and shape of the cavity of donors generated by the ligand, e.g. a protein. It is readily shown that a large cavity favours higher coordination number and larger cations. Hence Mn^{2+} and Fe^{2+} are expected to be found in six-coordinate octahedral sites, Fig. 19, while Zn^{2+} is favoured in smaller four-coordinate tetrahedral sites [1,2]. The comparison between the binding of EGTA as opposed to EDTA indicates how steric hindrance within the ligand frame inhibits binding to small cations relative to larger ones, Figs. 15 and 18. We note that as S is a larger atom than O or N, S-donors are not as able to supply high coordination number sites as N, O donors. Hence Mn^{2+} and Fe^{2+} prefer N, O-donors in six-coordinate sites while Zn^{2+} prefers four-coordinate N/S sites.

Table 8
Preferred symmetry of donors around metal ions[a] [1]

Metal ion	Preferred symmetry
Ca^{2+}	Indeterminate
Mg^{2+}	Octahedral
Mn^{2+}	Octahedral
Co^{2+}	Octahedral, tetrahedral
Ni^{2+}	Octahedral
Cu^{2+}	Tetrahedral
Cu^{+}	Linear
Zn^{2+}	Tetrahedral, octahedral

[a] N.B. These preferences do depend on donor atom sizes. Mg^{2+} and Ni^{2+} prefer N, O-donors in six-coordinate sites while Zn^{2+} prefers four-coordinate N/S sites.

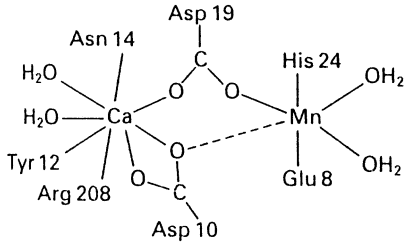
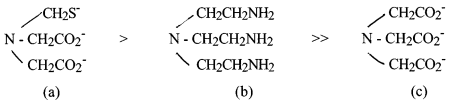


Fig. 19. The structure of concanavillin A showing the six-coordinate typical of Mn^{2+} (N→ O→ donors) and 7/8 coordination of Ca^{2+} O-donors only.

The second structural factor is the geometry desired by the metal atoms, electronic ligand-field effect, which at low coordination often favours Zn^{2+}, which has a very flexible geometric requirement for four or five neighbours, against the other cations. Ni^{2+} favours a six-coordinate octahedral disposition of ligands as do Mn^{2+} and Fe^{2+}, while Co^{2+} most closely mimics Zn^{2+}. We illustrate the effect of these preferences in Fig. 20 where we illustrate the binding of three ligands giving tetrahedral geometry and in order of binding strength for Zn^{2+} [23].



As will be seen the absolute binding strengths at pH 12 where there is no competition by H^+ for any ligating atom shows that Mg^{2+} prefers (c), but

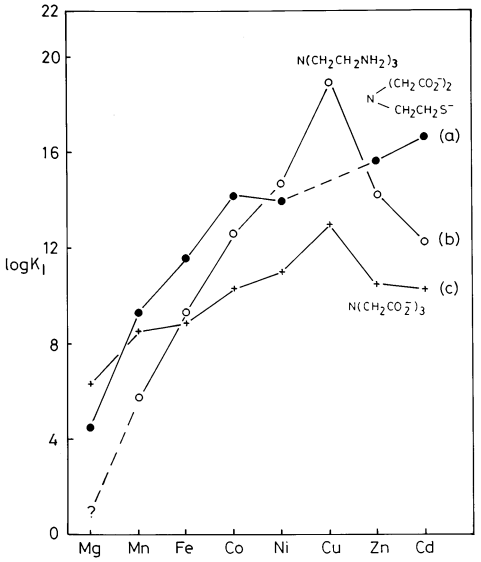


Fig. 20. A comparison of the stability constants for ligands using different donors: N→ only; N→, O→, S→; and N→, O→; where all the ligands fit a tetrahedron best. Data from [23].

increasingly the other cations prefer (a) or (b). There are some surprises however in that Ni^{2+} prefers (b) over (a). As stated earlier, full appreciation of the situation at pH 7 requires us to take into account competition from H^+. Let us suppose however that the N-bases are replaceable by imidazole of $pK_a = 7$. This is the case in enzymes and we know that imidazole binds as well as NH_3 to these metal ions. Thus the only correction we must apply for H^+ competition is a reduction of three log units to the binding constants, log K, for the $-S^-$ group of (a) with a $pK_a = 10.0$ as in enzymes. The interesting features are now that Mn^{2+} prefers (c); Fe^{2+} shows little preference between (a) (b) or (c); Co^{2+} prefers (b); Ni^{2+} prefers (b) strongly; Cu^+ (not shown) and Cd^{2+} prefer (a) while Zn^{2+} just prefers (b) but (a) and (b) are almost equal. Let us make one further adjustment to our thinking since enzymes generate centres with $4RS^-$; $3RS^-$,N; $2RS^-$, 2N and RS^-, 3N and 4N donors, respectively, where N is imidazole. We can estimate the effects of selection of any of these coordination conditions. The effect of changing from RS^- to N-donors will be to considerably decrease the stability of particularly Cu^{2+} (Cu^+) and Cd^{2+} at the beginning of the series. The decrease for Zn^{2+} and Co^{2+} is slight, and Ni^{2+} appears to prefer an all N-donor ligand in most circumstances. We shall see if all these conclusions arrived at from model complexes are reflected in proteins and enzymes (see Sections 4 and 5).

Finally we need to take into account spin-pairing electronic factors which can stabilise an ion. For example switch to low spin-states favours Fe^{2+} most strongly amongst the five ions under consideration (note also Co^{3+}) but we shall also need to consider spin states in metal ion clusters (see Section 6) where we consider redox enzymes. However there we are also interested in oxidation state changes for example to Cu^+ and Fe^{3+}

3.3. Metal clusters [24]

It is frequently observed that cations form clusters with certain anions, M_nX_m, in aqueous solution. The simplest statement we can make about their selective formation is that the lowest and highest oxidation states of atoms do not form clusters. In any period of the periodic table the extreme early groups of the series give rise to mononuclear cations, Groups 1 and 2, e.g. Na^+ and Mg^{2+}, while the later atoms of a period give mononuclear oxy-anions, Groups 16 and 17, e.g. SO_4^{2-} and ClO_4^-. From Group 3 to Group 15 there is observed a variety of multinuclear species with central atoms of heavier elements of high oxidation state or for ions of more than charge higher than two. There are exceptions in that Be^{2+} (Group 2), which is a very small cation and other divalent cations of high electron affinity later in the third period, e.g. Cu^{2+}, may also give multinuclear species. The conclusion must be that intermediate to high electron affinity of central atoms (small size, high oxidation state or charge >2 and considerable electron affinity) leads to the formation of clusters, e.g. Fe^{3+}, VO^{2+} and MoO_4^{2-} while lower or higher electron affinity does not. The major bridging groups are OH^-, O^{2-}, RS^- and S^{2-}. Obviously these units are relatively weak acid anions (strong conjugate bases) and

we know that strong acid anions (weak conjugate bases) do not bind positively. charged centres well. The overall conclusion is that low surface charge density on the central atom or on the periphery of an anion such as MO_4^- does not allow cluster formation. The stoichiometry and shape of the clusters is not easily described from this rule.

Some examples show the difficulty in predicting the nature of clusters. Fe^{3+} forms a $[FeO \cdot Fe]^{4+}$ cluster while central ions of higher charge and smaller size (and therefore higher electron affinity) such as V^{5+} and Mo^{6+} form much more extensive clusters based on octahedra. With S^{2-}, Fe^{3+} is able to form $[Fe_4S_4]^{4+}$ but here the Fe^{3+} is tetrahedral due to the large size of S^{2-} relative to Fe^{3+}. The problem of cluster formation of the kind [M–O–M] for Fe^{3+} and Mn^{3+} and of Fe_nS_mM for a variety of metal ions (M = Fe, Mo, V, Ni) is acute in proteins (see Section 6.2).

Now clusters also form using RS^- and carboxylate, $R \cdot CO_2^-$ ligands. With RS^- ligands clusters are obtained with metal ions such as Co^{2+} and Ni^{2+}. Once again we shall find corresponding clusters in proteins including those of Zn^{2+} and Cu^+ but not of other metal ions. However mixed $[Fe^{2+}/Fe^{3+}]_nS_m^{2-}$ and $[Fe^{3+}]S^{2-}$ clusters interact with RS^- ligands to form more extensive units. There are in fact a great variety of clusters in organisms and we shall consider that the cavity made by the protein is a decisive factor in cluster formation.

3.4. Summary of model equilibrium systems

Put together it would seem that an organism might be able to separate at equilibrium all divalent metal ions by a procedure much like that used by analytical chemists. The organism has the added advantage of feedback control over the amount of a given protein partner which is to be produced. Thus in a strong-binding partnership of a given metal ion, M, with a given protein, P, no appreciable excess of M or P is permitted (see Section 2 on the functioning of the genome). If this equilibrium condition held strictly all different MP proteins of any one of the different metals would have to be of closely the same degree of formation if they are to be useful, but this condition would only hold opposite very different free ion concentrations of each of the different metals, Fig. 2. The implication is that all the exchangeable binding sites for one metal must be of similar thermodynamic stability so that they are equally occupied. It is not difficult to see that such limitations on production of different kinds of ligands leads to a selective incorporation of elements in proteins, see below.

The selection principles described above will be shown to apply in cellular fluids in most cases but we have to note that a particular structural design of a ligand could lead to a cavity of very high affinity for any chosen metal ion with almost any set of donors. We must be prepared to uncover a few exceptions to predictions based on the above simple models.

If the above explanation of selection based on thermodynamic stability of model complexes proves on inspection to be inadequate for biological ligands (proteins) we can only fall back on kinetic explanations which use the following scheme.

STEP 1 Reversible thermodynamic selection generating MP complexes of (slightly) different stereochemistry for different metals.
STEP II Recognition of a given MP complex through its stereochemistry by an inserting system.
STEP III Insertion from MP into MP′ in an irreversible step, which may require energy.
STEP IV Deliberate bias of the competition between cations by removal of one of them into a separate compartment using a selective pump, i.e. using step II at some stage.

Given the complications of these distribution modes, the first task is to understand the selective binding of metal ions to the proteins, often not enzymes, which *equilibrate* with free ions. These proteins are largely pumps, carriers and transcription factors of necessity but it is required that the metal then generates a specific shape to the protein so that the combination is recognised. A more mobile protein is useful for selective recognition of metal ions and for faster exchange.

In Sections 4–8 we shall attempt to uncover the processes which achieve final selection in enzymes in equilibrium or non-equilibrium circumstances by first analysing the association of different metal ions with those enzymes which carry out acid–base reactions such as hydrolyses since a large number of them have been studied. It is easier to analyse the association of the metal ions with these enzymes than with redox enzymes since the properties of the metal ions which are important in the former, Lewis acid strengths, are so nicely graded along the Irving–Williams series which applies equally to their binding constants and their intrinsic catalytic ability as Lewis acids. The analysis of the use of different metal ions in redox enzymes is much more difficult since the additional idiosyncratic property — the relative stability of the complexes of different redox states of the metal ions — becomes very important. We describe them in Section 6.

3.5. Anion selection

The selective binding of anions is a very difficult topic since the simple electrostatic energy terms of spherical ions such as Na^+ and Ca^{2+} amongst cations are only applicable to mono-atomic anions such as Cl^- and I^-. Most anions are more complicated, e.g. NO_3^-, HPO_4^{2-}, CO_3^{2-} and a huge variety of organic anions. Here three factors dominate binding to organic surfaces: size, shape and hydrogen-bonding. To some degree hydrogen-bond strength can be estimated from the values of pK_a, acid dissociation constants, but selective fitting of shape and size is more difficult. Size fitting is easily illustrated for simple anions by the differences between the radii of halide anions (Å); F^- (1.30), Cl^- (1.85), Br (2.05) and I^- (2.35). Here differences are greater than between Na^+ and Cs^+. However proteins are known which can discriminate between SO_4^{2-} (3.1) and either MoO_4^{2-} and WO_4^{2-} (3.5). A more subtle discrimination is that between HPO_4^{2-} and SO_4^{2-} which have the same size. In this case H-bonding is of great importance. Shape fitting depends on short range repulsive as well as attractive forces so that since repulsion depends so highly on distance, $1/r^n$ where n is 6–10, very small changes in size give rise to major changes in binding. For a discussion of anion binding to model ligands, see Refs. [1,2].

3.6. Rates of exchange of metal ions [25]

Several factors affect the rate of exchange of a metal ion from a ligand complex in water.

1. The hydration of the ion;
2. The relaxation of the ligand;
3. The strength of interaction measured by the stability constant, K.

In the simplest case of equilibrium the value of K is related to exchange rates

$$K = \frac{\text{off rate (s)}}{\text{on rate (s)}}$$

The on-rates for ions, even of different charge types which have roughly the same diffusion rate, is dependent on the rate of water exchange [25] which follows the order

Na^+, K^+, Ca^{2+}, Cu^+, Mn^{2+}	10^{-9} s
Zn^{2+}, Cu^{2+}, Co^{2+}, Fe^{2+}	10^{-8} s
Mg^{2+}	10^{-6} s
Ni^{2+}	10^{-4} s
Fe^{3+}	10^{-3} s
Al^{3+}	10^{-3} s

The off rates can then be estimated from binding constants when in cells some approximate values are:

Na^+, K^+	10^{-9} s
Ca^{2+}, Mg^{2+}	10^{-3} s
Mn^{2+}, Fe^{2+}	10^{-2} s
Zn^{2+}, Co^{2+}	10^2 s
Cu^+	$>10^2$ s

Now if the on or off rates are restricted by (ligand) protein conformational changes then the rates can become extremely slow. for example the off rates of zinc from two proteins are:

Metallo-thionein	~ 10 s
Carbonic anhydrase	3 years

Note that metallo-thionein apoprotein is a random coil while apocarbonic anhydrase is a β-barrel, and metallo-thionein is a buffer or carrier where equilibrium exchange matters while carbonic anhydrase is an enzyme where generally it does not matter.

While this huge variation is found for zinc we consider it likely that virtually all Ca^{2+}, Mg^{2+}, Mn^{2+} and Fe^{2+}, enzymes exchange metal ions readily 10^{-3} s or at least relatively rapidly (≤ 10 s). By contrast Zn^{2+}, Ni^{2+}, Cu^{2+} and perhaps Co^{2+}

must be in slower exchange since they are most strongly held in complexes. This can lead effectively to irreversible binding while other metals are reversibly bound so that their activities are related to the bound ion concentration in their local environment. However in organisms we shall have to be careful in our descriptions since: (1) the metal ions and their proteins may be required to function in special environments, cytoplasmic, vesicular or extracellular; and (2) several of the metal ions can exist in different oxidation states — Mn, Fe, Co, Ni and Cu not Mg and Zn.

The exchange of simple anions from binding sites is generally fast especially for weak complexes with protonated organic bases. There are relatively few studies but fast exchange is found for various phosphate anions and we shall see that this makes them particularly suitable as messengers in living organisms.

4. Metal ion protein binding at equilibrium

4.1. Introduction

The way in which different metal ions are combined in different proteins has been analysed for nearly 50 years [26,27]. The suggestion, relating to binding of metal ions of a fixed charge type, divalent ions, based upon studies of model systems as analysed above, is that as the electron affinity of the metal ion increases, e.g. in the order of the 'a'-class to 'b'-class (hard to soft) for divalent metal ions (see Table 7) so there should be a switch in ligand binding preference from oxygen (O) donor centres (carboxylates, ethers, carbonyls, alcohols and so on) toward nitrogen (N) donors and finally toward sulfur (S) donors. At the same time, as the size of the ion decreases, lower coordination numbers should be favoured while lower coordination should also be found with larger donor atoms, $S > O$, N. In addition atomic overlap of orbitals should affect certain ions to favour specific symmetries. Generally anionic ligands bind metal ions better than neutral centres so that the weakest electron acceptor metal ions are expected to be bound by anionic O-donors since O-centres, e.g. carboxylic and phosphoric acids, generate anions more readily than N- or S-centres. In the broadest terms we now know from metallo-protein structure data that selection of divalent metal ions in organisms by many proteins follows these rules derived from inorganic coordination chemistry [26], Table 8. Thus we observe that at the extremes amongst divalent metal ions the weakest acceptors (hard ions), small divalent magnesium ions (six-coordinate) and large calcium ions (seven- or eight-coordinate), bind only through anionic and neutral O-donors, while the strongest acceptors (soft ions) cadmium ions (four-coordinate) and mercury ions (two-coordinate) bind only through S-donors. However, even here there is one striking exception: magnesium ions bind to N-donors (five coordinate) in chlorophyll containing proteins. We consider this incorporation later as a kinetic trap (see Fig. 32). Notice that we might conclude that the majority of metal ion protein combinations appear to obey distribution rules based on thermodynamic equilibria, as they parallel generalisations from inorganic model studies which are based on

such equilibria. If this is so then there must be a stationary free divalent metal ion concentration, different for each metal ion in a cell compartment of a given organic chemical composition (see Fig. 2).

Now if the above generalisations, based on model equilibrium thermodynamic data, hold then the selection by proteins along the Irving–Williams series, again of divalent metal ions, that is along the series of increasing electron affinity and decreasing ion size (Table 7), and for single metal ion sites,

$$Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$$

should show a changing tendency to demonstrate first largely O-donor binding at Mn^{2+}, then more N-donor binding tendency through Fe^{2+} to Ni^{2+} and finally at Cu^{2+} and Zn^{2+} mostly N- and S-donor combinations alone while the coordination number falls along the whole series. The considerable number of structures now available show that to a first approximation these general statements are often found [1]: Mn^{2+} (five- or six-coordinate) is bound by O-donor ligands plus one, but usually not more, N-donor (histidine) ligands, while particular Cu^{2+} (Cu^{+}) and Zn^{2+} are found bound mostly frequently to thiolate groups and to histidines and are then at most four coordinate. Fe^{2+}, Co^{2+} and Ni^{2+} bind most often to an equal mixture of N- and O-donors when in single metal ion sites and are often five- or six-coordinate. We will refer to clusters later, Section 5.14. Note that there is some selection of the symmetry of packing of ligands around the metal ions and we know that the energies of favoured symmetries around different ions must be considered, Section 3.2, but this does not appear to be a very important term in the proteins under discussion since there is little match between preferred symmetry Table 8 and that observed, Table 9. The reason for this failure is that preferred geometries are frequently the most inert. More detailed study of all data on both single divalent metal ion sites (see Table 15) and two (or more) metal sites (see Table 16) in enzymes, however does not give such a clear picture. First, turning to extremely unusual patterns of chosen donor atoms and coordination numbers, iron,

Table 9
Ligand field geometric preferences and observed site geometries in proteins [1,2]

	Coordination in models	Coordination in proteins	Ligand donors
Ca^{2+}	7/9-coordination	7/8-coordination	O
Mg^{2+}	6-coordinate (octahedral)	six-coordinate (octahedral)	O
Mn^{2+}	6/7-coordination (octahedral)	5/6-coordinate irregular	O, ((N))
Fe^{2+}	6-coordinate (octahedral)	Often 6-coordinate irregular	O, (N) see clusters (S)
Co^{2+}	6-coordinate (octahedral) 5-coordination	5-coordinate variable[a]	O, N
Ni^{2+}	6 coordination (octahedral)	4/5/6 coordination[a] irregular	O, N
Cu^{2+}	4/5 coordinate (tetragonal)	3/4 coordination irregular	N, S
Zn^{2+}	4/5/6 coordination tetrahedral(4) octahedral(6)	4/5 coordination irregular tetrahedral	N, S, O

[a] Very few examples.

Table 10
Metal ions in biological compartments [1,2]

Metal ion	Compartment	Proteins involved
Mg^{2+}	Cytoplasm	Pump
Ca^{2+}	Endoplasmic reticulum	Pump/calreticulum exchangers
Mn^{2+}	Golgi	Pump
	Mitochondria	Pump
Co^{2+}	Cytoplasm	Pump
Ni^{2+}	Plant vacuoles	Pump (store)
Cu^{2+}	Golgi/endoplasmic reticulum/extracellular	Pump/many
Zn^{2+}	Vesicles	Pump/Many
Na^{+}	Extracellular	Pump exchanger

cobalt and nickel are all found in very similar N-donor rings of porphyrin-derived chelates in heme, corrin and F-430, respectively, and all are bound quite specifically. There is no obvious basis for this selection. We set these cases aside with that of magnesium chlorophyll mentioned above, returning to a discussion of them in Section 8. Second there are some observations on direct protein ligation through amino acid side-chains that do not appear to fit at all well with expectation. How can divalent nickel and zinc be so clearly separated from one another in different enzymes but often with similar donor atoms and geometries (see Table 16)? Third why does leucine amino peptidase bind zinc and not manganese, ferrous or cobalt ions since all ligands are O-donors (see Table 16)? These facts must lead us to seek considerations quite different from those based on comparison with the absolute thermodynamic binding data used in the above discussion and derived from equilibrium (reversible) model studies. Let us allow that there is often a direct and simple thermodynamic selection of metal ions particularly for weakly binding ions such as Mg^{2+}, Ca^{2+}, Mn^{2+} and Fe^{2+} but in order to explain the more idiosyncratic associations mentioned above we need to ask in what other ways could a particular association of a given metal ion and a given protein arise.

As stressed in Table 10, one way in which to bias the thermodynamic equilibria is to establish compartments within which one metal ion rather than any other is concentrated together with particular proteins. Thus we must direct our attention to the way positioning of an element (and a protein) can be biased by creating, using energy, a preferred transfer of a given metal ion across a membrane. A fine example is the pumping of Ca^{2+} into many vesicles from a cytoplasmic concentration of 10^{-7} up to 10^{-3} M. Mg^{2+} is not pumped into these vesicles so when ATP is also pumped into vesicles as in the adrenal gland it is bound in part with Ca^{2+}, not Mg^{2+}, as in the cytoplasm [2]. Any metal ion can be pumped as a bare cation or as a complex in reversible equilibrium with free M^{2+} but, as we have made clear in Section 2, it is the selection by the pump edges and not the gradient across the membrane that is based upon thermodynamically reversible equilibrium constants, Fig. 21. A membrane is not essential for a particular kinetically controlled metal/protein attachment, since specific insertion can be used to force any given ion

into an unusual site using an irreversible route of transfer within a given phase, Fig. 3. Notice that some form of thermodynamic selection of ions is however necessary before pumping or insertion can occur in a kinetic selective manner. After these steps further selection can be based on the molecular shape of the bound complex, as we noted in the cases of transfer of small covalent compounds of H, C, N, O, S and P, or change of oxidation state, which is a more selective property than binding of a given charge type.

Clearly, if we are to understand the ways in which metal ions and their proteins are brought together, we must separate three distinct processes for a given charge on the ion:

1. Equilibrium binding in a given compartment, taking each compartment separately.
2. They way in which the metal ion and protein are transferred into a given compartment.
3. Insertion of a metal ion into a particular protein in a given compartment.

As stated earlier, in most cases there is observed additionally and especially for strong binding systems a direct link between limited final amounts of proteins for a limited final amount of a given metal so that production of a given protein must be related to the concentration of the one metal ion it holds. This requires a set of feedback relationships between the controlled metal ion concentrations and regulated protein production on a selective basis. We turn first to the equilibria in given compartments.

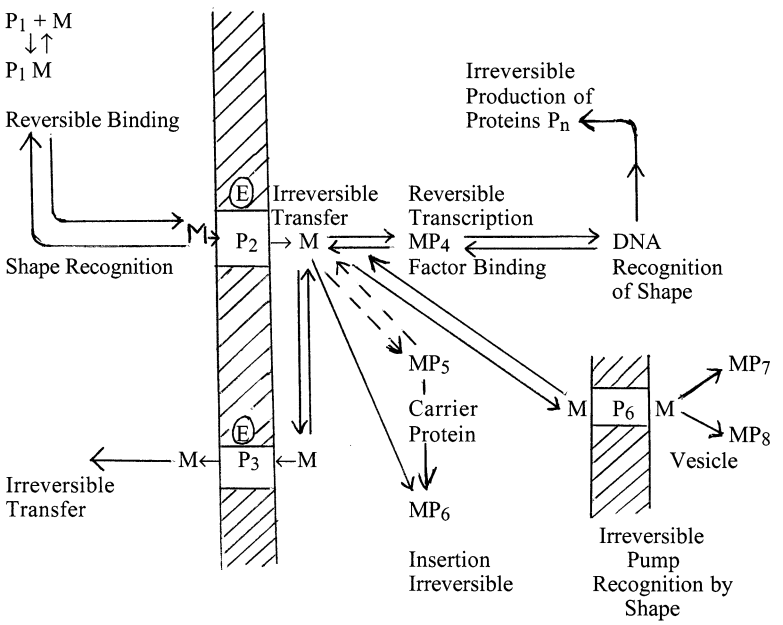


Fig. 21. An elaboration of Fig. 3 showing the importance of binding strengths, reversibility and shape for the distribution of elements.

The first question to ask is which types of proteins in cells have to be in equilibrium with the free metal ion concentration. It is obvious that transcription factors which help to control the free ion concentration must be in equilibrium between bound and unbound forms. Again a carrier of an ion may have to both pick up and release ions fairly rapidly so it too may well be in equilibrium with the concentration of the free metal ion but this is not essential. As stated earlier a pump or an exchanger is a carrier through a membrane and during transfer the binding constant of the pump to the metal ion is changed. Binding on each side of the pump (exchanger) then relates to the free concentration of ions in the phase in contact with it. The pump ceases to operate below the binding constant level of the ion. Hence the ions on the side of the pump facing the phase containing transcription factors and carriers, that is the cytoplasm, may well be in equilibrium binding to the ion concentration with all three types of protein. The face of the pump toward the outside will be in equilibrium with the extracellular fluids. The final destination of the ion in the cytoplasm may be an enzyme or a structure. In these final situations equilibration may well not be achieved or indeed desirable, Fig. 21. However if it is required that an enzyme is switched on and off at a certain level of the metal ion concentration then the metal-enzyme will also equilibrate with the free ion concentration. Equilibrium is clearly not necessarily desirable if the metal-enzyme is transferred to another compartment which has an ion concentration very different from the solution in which it was initially found; for example to extracellular fluids or some vesicles. Here loss of a metal could be destructive of an enzyme. We conclude that we should examine binding to pumps, carriers and transcription factors which are in equilibrium in the cytoplasm before we analyse proteins, menzymes (or carriers), which may or may not equilibrate in the cytoplasm, in vesicles or extracellular fluids. Do not forget that the transcription factors see to it proteins are produced commensurate with the free metal ion present so that neither is in excess to any marked degree.

4.2. Metal ions and proteins in equilibrium

As stated above in a system of equilibrium binding with fast exchange of one metal ion between several proteins then in order that the different proteins should be roughly equally occupied they must have approximately equal binding constants. It follows that equilibrium binding to pumps, carriers and transcription factors for a given metal ion all in the same phase must have nearly equal binding constants. At the same time a factor of some ten to one hundred must separate the probability of the particular metal binding to a protein compared to those of all other metal ions if the system is to be selective. This statement does permit equivalently strong binding of two or more metal ions but only if, due to a variety of factors stated in Section 3, they bind to very different ligands, e.g. Cu^{2+} and Ca^{2+} could equilibrate with an N/S donor and a O/O donor where all four components are in equal concentration. The Cu^{2+} will then be bound to the N/S centre while the Ca^{2+} will be bound to the O/O centre to almost 100% even though both ions bind equally with $K > 10^7$ to this O/O centre. Cu^{2+} is effectively sequestered and prevented from competing with Ca^{2+}.

Table 11
Cytoplasmic and extracellular proteins at equilibria

Metal ion	Protein	Estimated equilibrium K	Ligands	Ref.
Ca^{2+}	Calmodulin	10^7 feedback to pump	O	[30]
	Calbindin	10^7 carrier	O	[30]
	ATP-ase pump	10^7	O	[30]
Mg^{2+}	MgATP	$\sim 10^3$ feedback to pump	O	[1,2]
Mn^{2+}	Fur	10^6 transcription factor	O(N)	[31]
Fe^{2+}	Fur	10^7 transcription factor	O/N	[31–33]
	Aconitase	10^7 transcription factor	Fe_3S_4	[34]
	Fnr	10^7 transcription factor	Fe_3S_4	[34,35]
	DCT/N ramp 2 [a]	(?) carrier	O/N	[82]
Co^{2+}	(?)	10^8 carrier/pump	N(S)?	[36,37]
Ni^{2+}		(?) 10^9 carrier/pump	N?	[38,39]
Cu^{+}	Ace-link	$\gg 10^{12}$ transcription factor	RS^- (cluster)	[2]
	ATP-ase pump	$\gg 10^{12}$	RS^-	[40]
	Metallo-thionein	$\gg 10^{12}$ carrier(?)	RS^-	[2]
Zn^{2+}	Zinc-fingers	$>10^{11}$ transcription factor	RS^- (imid)	[41,44]
	ATP-ase pump	$>10^{11}$	(N)RS^- (N)	[43]
	Metallo-thionein	10^{11} carrier(?)	RS^-	[1,2,42]
	Zur	10^{11} (?) transcription factor	N(RS^-)	[32]
Fe^{3+}	Transferrin	10^{21} carrier	ϕO^-/N	[2]
Zn^{2+}	Albumin	10^{12} carrier	N	[2]
Cu^{2+}	Albumin	$>10^{15}$ carrier	N/N^-	[2]
Ca^{2+}	Phosphoprotein	10^3 carrier	O	[2]

[a] The examples below the dashed line are of extracellular carriers.

4.3. Equilibration: pumps, transcription factors and some carriers

Table 11 lists some of the known carriers and pumps, and some of the known transcription factors associated with metal ions [27–29]. We shall assume at first that they all equilibrate. The implication is that, inside one compartment, MP_2, MP_3, MP_6 pumps, and MP_4, MP_5, MP_6 carriers and transcription factors, Fig. 21, have virtually identical binding constants to the metal ions and this must be true for all other functional activities which are at equilibrium. We know that this situation holds for Mg^{2+} and Ca^{2+}, which are held close to 10^{-3} and 10^{-7} M, respectively, in the cytoplasm of all cells. At these levels their binding systems are known to be very selective, not just between themselves but against all other metal ions. They are based on O-donor ligands which are anionic [30]. The question arises as to how these sites avoid binding to other metal ions forming complexes which have higher absolute binding constants. When we examine the pumps, carriers and transcription factors for other metal ions, e.g. Zn^{2+}, then we find that their very strong binding uses some N- or S-donors, Table 12, which do not bind Ca^{2+} or Mg^{2+} at equilibrium. Moreover, these and similar donors reduce the concentrations of these competing elements in the cytoplasm, Fig. 2, to below the value at which they could still bind to the O-donor sites for Ca^{2+} or Mg^{2+}, Figs. 17 and 18. For example the

bonding to ATP which is free at about 10^{-3} M in cell cytoplasm cannot be to any of the ions Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Ca^{2+} or Zn^{2+} due to the low concentrations of these ions in cells. The proteins induced by transcription factors together with equilibrium feedback to DNA ensure that no further amounts of these metal ions or their proteins are incorporated or synthesised after their effectively saturated binding is reached. Now we observe that following our findings from the analysis of the binding constants in model complexes the binding strengths to transcription and carriers of different kinds, i.e. which use different donor coordinate centres, follows expectation taking into account the size, the electron affinity (Table 7), and the stereochemical preferences of metal ions (Table 8). Note that the selection even applies to that of Fe^{2+} by Fe_3S_4 centres (see Section 6.2).

One implication of these findings is that the free ion levels of the elements are in the inverse order of their *binding* constants to ligand specifically 'designed' to bind them so that in fact the potential ability to bind to the ligands 'designed' for other metals of weaker strength is reduced. As Table 16 shows some proteins bind to more than one different metal ion and have more than one specific site (see Section 4.4). Note that there are no *outward pumps* from the cytoplasm for free Fe^{2+} or Mg^{2+} which are not discharged out of cells or into vesicles, while there are pumps for Ca^{2+}, Mn^{2+}, Ni^{2+}, Cu^{+} and Zn^{2+}.

4.4. Two examples: Fur and Zur [31–33]

Fur and Zur are the two bacterial uptake (u) regulatory (r) proteins for iron (Fe) and zinc (Zn). In fact Fur also has a zinc binding site. The proteins are transcription factors of similar sequence and presumably structure. The two binding sites exchange with free zinc (site A) and free iron (site B) selectively. The selection is achieved by the chemical differences in binding site, strengths of binding and the limitations on free metal ions by buffering and/or pumping, and limitations on protein production by regulation of DNA. The site for zinc is tetrahedral using two RS^{-} and probably one histidine (N) donor giving a binding constant of about 10^{11}

Table 12
Zinc binding sites in the cytoplasm[a]

Proteins	Donor groups
Single stranded DNA/RNA	Cys. His. Cys. Cys.
Binding proteins	Cys. Cys. His. Cys.
Double stranded DNA binding	Cys. Cys. Cys. Cys.
Proteins (zinc fingers)	Cys. Cys. His. His.
(Ten classes)	
Metallo-thionein	Cys. Cys. Cys. Cys.
Alcohol dehydrogenase	
Aspartate transcarbamylase	Cys. Cys. Cys. Cys.
Protein kinase-C	

[a] See Ref. [29].

which is related to that for metallo-thionein, the zinc buffer [2]. The iron (Fe^{2+}) site is octahedral with one or two histidines and one or two carboxylates. Its binding constant to the protein is around 10^7 [31]. It is easy to see that although the first site (A) may not be a very good site for Fe^{2+} the second site is (see Figs. 17 and 18). In the presence of limiting concentrations of Zn^{2+} (around 10^{-11} M) and Fe^{2+} (around 10^{-7} M) the two sites will be occupied by zinc and iron separately. The Fur Fe-site may also be occupied preferentially by Mn^{2+} in the relative absence of Fe^{2+} while the Zur-site is preferentially occupied by Cd^{2+}, a poison.

4.5. Summary of equilibrium binding to proteins

It is now clear that it is the *effective* binding constant under the *buffering* and *pumping* conditions of the cytoplasm which defines the true competition for any ligand. We cannot give exact numbers but we consider that the concentrations in the cytoplasm of many cells could be as different as shown in Fig. 2, where it is seen that more than ten-fold differences frequently exist:

Cu^{+}	Zn^{2+}	Ni^{2+}	Co^{2+}	Fe^{2+}	Mn^{2+}	Mg^{2+}	Ca^{2+}
10^{-15} M	10^{-11} M	10^{-9} M	10^{-8} M	10^{-7} M	10^{-6} M	10^{-3} M	10^{-7} M

Several conclusions can be drawn from this analysis of equilibrium binding to proteins, most of which are not enzymes.

1. In all cases the level of production of an apoprotein and the level of concentration of a metal ion are controlled by feedback so that only very low concentrations of free metal ions (except for Mg^{2+}) or free apoproteins (except for Mg^{2+} binding proteins) are allowed.
2. A cell operates opportunistically at levels of free metal ions roughly commensurate with their greatest ability to bind to organic molecules especially proteins and given the side-chains which these proteins have. Different metal ions may be present at similar total amount present and similar free ion concentrations but they are then bound to different proteins through different donor groups.
3. It is not easy to achieve clear-cut distinctions between very similar ions. Examples are the binding to Fur of both Mn^{2+} and Fe^{2+} and the binding to metallo-thionein of Zn^{2+}, Cd^{2+} and Cu^{+}. In each case the different structures imposed on the protein by the specific metal ion allows recognition at DNA or at a membrane where there follows irreversible activity, expression or transport, specific to the particular metal ion, Fig. 21.
4. We must also remember that the concentration of all ions in the cytoplasm or vesicles can be controlled by pumping as well as buffering. We know that Ca^{2+} is held at $< 10^{-7}$ M in the cytoplasm against external and vesicular concentrations of 10^{-3} M. It is also known that other metal ions are pumped out of cells and into vesicles so that generally they are all kinetically managed in the cytoplasm. Certainly Zn^{2+} can be as high as 10^{-4} M in some vesicles; Ni^{2+} is high (10^{-5}M) in some plant vacuoles, and Mn^{2+} is $> 10^{-6}$ M in the Golgi and quite possibly so in both mitochondria and chloroplasts. Thus binding to

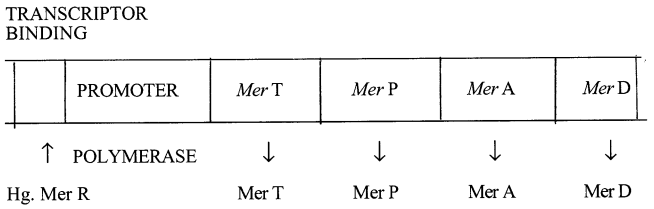


Fig. 22. The sequence of genes for the production of mercury binding proteins in *E. coli*.

exchangeable sites is governed by the free concentration locally multiplied by the binding constant.

5. The way metal ions and proteins are put together only obeys equilibrium conditions in a given compartment to a limited extent. The selection removes the weaker Lewis acids by O-donor binding and stronger Lewis acids increasingly by N- then S-donor binding. In Section 4.6 we examine the irreversible steps which lead to distribution not always closely related to the above, especially in enzymes.

Now that we have described these metal/protein bindings which must come to equilibrium we can go forward to the examination of metallo-enzymes. Increasing the discussion involves the homeostasis of conditions in organisms in their dynamic states where there is kinetic rather than thermodynamic control.

4.6. Metal ions, their proteins and gene structure

Now the demand to make suitable proteins according to the amount of metal ion in a system means that there has to be a set of transcription factors which bind adjacent to promoter regions (Fig. 1) for the specific metal-proteins. In the simplest case there is just one transcription factor for all the gene products relating to a given metal ion. The case of mercury is the induced genetic system shown in Fig. 22 [29]. The proteins Mer T, P, A, and D are required for binding and removal of mercury from the cell. They are induced only when protein Mer R is bound to mercury. Mer R (unbound) may be constitutively expressed and always attached to the transcriptor binding site.

For other metal ions genes are often under multiple transcription factor site controls so that the gene structure for say zinc proteins is very complicated.

4.7. Homeostasis of protons

We shall describe in a series of sections the way in which a cell maintains the metal ion distributions which we observed in the previous section. We start with an account of the proton homeostasis the proton is the simplest cation and often the pH affects almost all bindings of ions and molecules in cells and is in equilibrium in each aqueous compartment. Note that the cell is a *dynamic system* but as far as possible we shall describe steady states by equilibria in compartments which we realise is not thoroughly satisfactory. Homeostasis involves a flow of energy and material together so that we stress again that it is not the same as buffering.

The proton, H^+ is held at a fixed concentration in many biological compartments. Some features are listed in Table 13. We can ask what is the evolutionary advantage of pH 7.3 in the cytoplasm? Later developments in evolution gave rise to different pH values in different compartments and again we can ask why is this so? Now several common anions of the cytoplasm have pK_a values close to 7.0. The value for phosphate esters and anhydrides is 6.5 to 6.8, the value for HCO_3^- is ~ 6.5 and the value for imidazole is 7.5. Hence, given the nature of biological solutions in the cytoplasm at pH 7.0, they are quite strongly buffered by the several species which are present there. It is worth noting too that a pH close to 7.0 is such that both H^+ and OH^- are as low as they can be so that simple general acid/base hydrolysis by H_2O is reduced to a compromised minimum. Clearly the preservation of biological polymers from degradation is a critical consideration except in digestion. A pH of around 7.0 is then clearly advantageous for the development of selective catalysis by enzymes held in intracellular or extracellular animal fluids at pH 7.3. That the plant extracellular fluids are at pH 5.5 is curious and not readily explicable. Is cellulose open to alkaline hydrolysis but not to acid hydrolysis when pH 5.5 gives protection? A pH close to 5.0 in lysosomes clearly promotes protein hydrolysis and metal ion release.

While it is true that buffering close to pH 7 is in part an automatic consequence of the nature of biological polymers it is also clear that the pH in a cell is more exactly controlled. Now the difficulty in resolving the required mode of feedback to hold the value in homeostasis steady lies in the innumerable activities in a cell which are pH dependent. If a cell pH is drifting locally in a compartment toward say more alkaline values the cell undoubtedly compensates by producing acids. Only in this way can we explain that each compartment shown in Table 13 is held at a fixed and different pH. In fact there are well-known proton ATP-ases which pump H^+ into some compartments and out of others. These proteins are constitutively expressed in many cells of different kinds. The proton gradients are often also exchange-coupled to the movement across membranes of not only ions such as Na^+ and Ca^{2+} but also molecules such as amino acids. There is then a web of acid/base connections between $[H^+]$ and many other concentrations so that while we know cells control pH locally in compartments the explanation of the control is complicated. $[H^+]$ is only maintained at equilibrium selectively in particular compartments and this is true also for most metal ions.

Table 13
The pH of some biological fluids

Fluid	pH
Cell cytoplasm	7.3
Animal extracellular	7.3
Plant extracellular	5.5
Some lysomal vesicles	5.5
Digestive systems of animals	1.0
Thylakoid (in light)	3.5
The sea	8.5

The hydrogen atom, $H^{\bullet}$, not the hydrogen ion, H^+, is distributed by various redox carrier co-enzymes such as NAD, quinones, and flavins as well as by numerous organic substrates (see Section 6). The conversion of bound hydrogen to protons (in proton gradients) and then to ATP is a major source of energy for most organisms [19]. However, this transport of bound hydrogen by this type of organic molecule does not equilibrate with the proton concentration. (Even the proton gradient may be localised as is the calcium gradient as activity changes [19]). Hence the various co-enzymes and substrates carrying H can work simultaneously in a cell at different redox potentials. By way of contrast the co-enzymes glutathione in the cytoplasm and ascorbic acid in the extracellular fluids appear to exercise control over the redox state of many oxidation-reduction reaction pairs in equilibrium with the pH. The control of compartmental redox potential by bound H/H^+ is then as important as the control of pH and is linked to a range of reactions. It would appear that the redox potential of the cytoplasm cannot go higher than the glutathione potential, -0.25 V, but the redox potential in many vesicles (and organelles?) is very differently poised. Note that generally $H^+/H^{\bullet}/H_2$ is not in equilibrium anywhere in a cell.

The local connection between redox reactions and pH gradients across membranes of organelles is the basis of one major mode of energy capture by cells, and is the subject of quite separate discussion, [1–3]. Once it is used to make ATP it is only partially reversible.

Finally as well as pH gradients movement of uncompensated charge across membranes creates potentials. Potentials or fields across membranes interact quite generally with all charges in a non-specific way but channels within membranes have selectivity so that potentials created by one ion, e.g. H^+ or Na^+, can be used, coupled, to the movement of another selected ion, e.g. Ca^{2+} in a process of exchange. Within any compartment the potential equilibrates. *Protein exchangers* in membranes could operate close to overall equilibrium and frequently the proton is involved. However energy coupling to pumping is irreversible. Once again proton homeostasis is linked to that of many other elements and small molecules.

When we were describing the effective binding constants of metal ions and model ligands we pointed out that it was very necessary to define the pH of the solutions since the proton frequently is a competing cation. In the next sections we shall keep this in mind while we describe the homeostasis of elements other than hydrogen.

5. Metallo-enzymes [45]

5.1. Introduction

The question we raise next concerns the selection of metal ions used in metallo-enzymes. The problem now includes aspects which are very different from those we discussed in the analysis of selection by metal ion carriers, pumps, exchangers and transcription factors in that we looked upon most of them as being always in equilibrium with free ions in a particular compartment, say the cytoplasm, Section

Table 14
Metal ions in biological compartments

Metal ion	Compartment	Protein comparison	Selection
Mg^{2+}	Cytoplasm	Kinases	Thermodynamic
Ca^{2+}	Endoplasmic reticulum	Pump/calreticulum	Kinetic
Mn^{2+}	Golgi	Pump/glycosidase	Kinetic
	Mitochondria	Pump/dehydrogenases	Kinetic
Co^{2+}	Cytoplasm	Pump/methionine peptidase	Kinetic
Ni^{2+}	Plant vacuoles	Pump/none (store)	Kinetic
Cu^{2+}	Golgi/endoplasmic reticulum/extracellular	Pump/many enzymes	Kinetic
Zn^{2+}	Vesicles	Pump/many enzymes	Kinetic

4. Enzymes are distributed more widely in many compartments including the extracellular fluids, Table 14, but since there they must not lose activity they must not equilibrate with free metal ions. We pointed out earlier that the problem is the irreversible transfer of the metal-proteins across membranes into solutions of different free metal ion concentrations. There is the further observation that in some compartments it seems to be necessary for a given metal to be irreversibly *inserted* into a particular protein in the solution of origin so that a metallo-enzyme as an intact unit may be kept in one compartment as a non-exchanging unit, Fig. 23. We gave the obvious examples of heme, vitamin B_{12} and F-430 dependent enzymes. In this section we shall be concerned only with direct binding to side-chains of protein amino acids where metal ions can also be in slow exchange. Even so many enzymes of metal ions such as Mg^{2+}, Mn^{2+} and Fe^{2+} do inevitably exchange metal ions readily since they do not bind sufficiently strongly. They can be grouped with equilibrated exchanging ions.

Before we turn to the analysis of the enzymes we have to be aware of the difficulty in knowing for certain in some cases which metal ion is the one that is

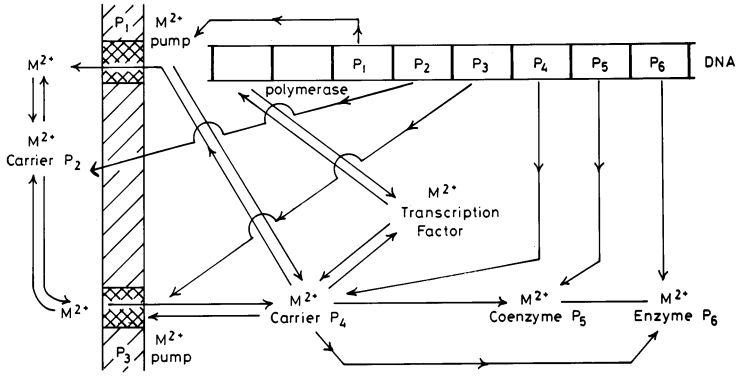


Fig. 23. The different transfer modes of metallo-proteins which allows irreversible selective combination in a specific compartment of a cell, on left, as well as equilibrated combination.

held and acts in vivo [46]. We cannot rely on the fact that a particular metal ion is functionally the best at equal concentration as shown by in vitro tests since we have just stressed that competition for a site in a cell is biased by the production of other proteins which control free ion concentrations at very different levels for different metals, Fig. 24. We cannot rely either on the metal ion being correctly identified by crystallographic analysis since all too often the metal seen in the crystals is that added to the protein or apoprotein in excess (1mM) by the experimenter during crystallisation. Hence we shall state in this article the most probable functional metal using knowledge of a binding site from a combination of information from isolation data, structures and the concentration of the ions concerned in a cell. Given the uncertainties in parts of this study we ask first which are the metal ions that we might reasonably expect to have been selected to act as catalysts in enzymes based on our knowledge of the catalytic activities of the ions in model reactions (see Section 3.6).

5.2. Kinetic strengths of metal ions in model systems [1,2]

The ability of a metal ion to act as a catalyst is undoubtedly related to (1) its ability to bind to a susceptible small molecule and (2) its ability to act as an electron withdrawing group. Since in the series of ions

$$Mg^{2+} < Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Zn^{2+} < Cu^{2+}$$

both abilities rise in the same order, the Irving–Williams order (see Table 7) overall catalysis of hydrolysis reactions in model systems follows this order no matter what molecule is attacked. This information leads to the expectation that in metal replacement experiments in metallo-enzymes rates will follow the same order as in

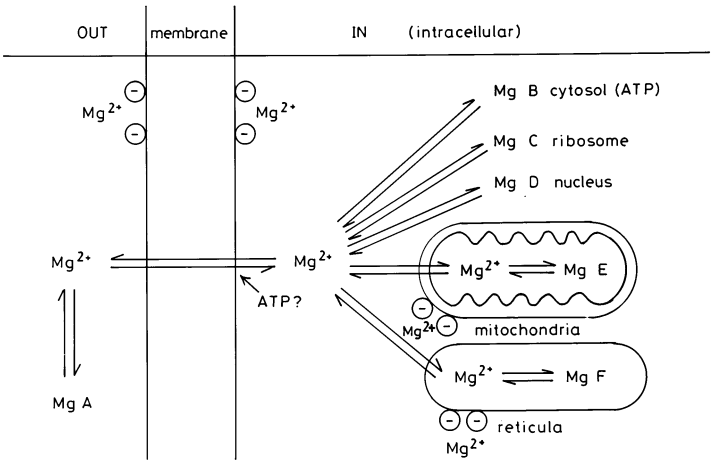


Fig. 24. The complicated connection of Mg^{2+} to many proteins and enzymes and to ATP, all in equilibrium, with binding constants around $10^{3.5}$ M^{-1}. Note virtually all the proteins are in the cytoplasm, see Section 5.5.

Table 15
Single metal ion sites in enzymes

Metal ion	Enzyme	Site	Donors	Ref.
Mg^{2+}	Enolase (see Table 17)	Octahedral	O	[1,47]
Mn^{2+}	Glycosidase (see Table 23)	Octahedral?	O?N	[48]
Fe^{2+}	Oxidases (see Table 26)	Octahedral	N, O	[49]
Co^{2+}	Not known (see Table 16)	–	–	[50]
Ni^{2+}	Not known (see Table 16)	–	–	[51]
Cu	Azurin	Trigonal (tetrahedral)	S, N	[1,2]
Zn	Carbonic anhydrase (see Table 19)	Tetrahedral (5-coordinate)	3/4 N, O N/S	[52–56]

the model series. However, this is never observed since a protein does not pick up all the metal ions in exactly the same way and the high affinity sites, for say copper, are often not those at the active site. Again a protein may well not need the most powerful Lewis acid, that is it can be that useful purpose is served employing Mg^{2+} rather than Zn^{2+}. Thus we find some enzymes which act best using Mg^{2+} and others which use Mn^{2+}, or Co^{2+},or Ni^{2+}, or Zn^{2+}. It may be that selection is based in part on rates of exchange since some of the metal ions could act as a self-controls over enzyme activity through exchange. Here a weaker Lewis acid cation, e.g. Mg^{2+}, can be functionally more valuable since it exchanges more readily. Finally the metal used may be a remnant of choices made very early on in the evolution of life, and which were kept unchanged although metal ion availability is now different. Some uses of single metal ions in enzymes are given in Table 15.

A very interesting development of model catalyses was introduced by Sargeson who showed that in the case of many substrates simultaneous attack by two (or more) metal ions acting differently brought about the most effective catalysis. Thus once again the best Lewis acid was not required even for a quite difficult hydrolysis. Of the two metal ions, the one which was the poorer Lewis acid, acts to assist the binding and appropriate orientation of the substrate while the second acts strongly as the attacking acid. One of the two, particularly the weaker, could also exchange and act as a control. We shall see that multiple metal ion sites using a pair of the same or two different metal ions are often found in enzymes, Table 16. The selection of metal ions in an enzyme is then complicated.

In the context of the above a further outstanding initial problem is to distinguish what is effectively reversible from irreversible binding within the biological context. This is an important consideration since in effectively reversible systems we can still analyse many properties in terms of equilibria. The difficulty lies in the time over which exchange is usefully analysed. Metal ions such as Na^{+}, K^{+} and Ca^{2+} exchange rapidly, faster than 10^3 per s, even when they have binding constants $> 10^6$. Since biological response times generally are $> 10^{-3}$ s these ions are in effective equilibrium exchange. Metal ions such as Cu^{2+} and Ni^{2+} usually exchange extremely slowly, especially since they have binding constants $> 10^{10}$. Reversibility has to be seen now in terms of their relative time of activity. Calcium

Table 16
Two-metal ion centres in enzymes

	Sequence								Enzyme	Compartment[a]	Ref.
	243	260	271	347	380	400					
Mn^{2+}	H	D	D	H	E	E			Arginase	C	[48]
									Proline amino-peptidase (see clusters)		
Fe^{2+}	90	92	118						Calcineurin see Zn^{2+} (see clusters)	C	[49]
	D	H	D								
Co^{2+}	D	D	H	E	H				Methionine amino-peptidase	C	[50]
	134	136	217	246	273	360			Urease	C	[51]
Ni^{2+}	H	H	K(O)	H	H	D					
	50	52	54	110	141	187					
Zn^{2+}	H	H	D	H	C/H	H			β-Lactamase	E	[53]
	1	14	55	69	118	122	128	146	Phospho-lipase C	E	[55]
	W	H	S	H	H	D	H	E			
	250	255	273	332	334				Leucine amino-peptidase	C	[52]
	K(O)	D	D	D	D						
	118	150	199	281					Calcineurin	C	[49]
	D	Asn(O)	H	H							
	51	102	155	322	369	370			Alkaline phosphatase	E	[57]
	D	S(O)	S(O)	E	D	H					
		Mg	Mg	Zn	Zn						
Cu^{2+}	H	H	H	H	H	H			Oxidases	E	[1,2]

[a] C = cytoplasm; E = extracellular.

concentration changes for example are used to switch on and off changes in cells in $\sim 10^{-3}$ s, e.g. mechanical contraction. Zinc and copper level changes are related to slow processes, e.g. development, relayed through transcription factors when exchange with a time constant of $\sim 10^{3}$ s may well be functionally significant. The situation parallels that of time constants of associated organic chemical signals, for example acetyl choline (action time 10^{-3} s) is related to the actions of Ca^{2+} in transmitter release steps, while sterols (action time $> 10^{3}$ s) are related to the actions of zinc in hormonal binding and response. Both reactions can be described as reversible and analyses of equilibria are then rewarding. For equally fast on-rates of binding, say 10^{-9} s at the fastest, this means that binding differences in the ranges 10^{6} M for Ca^{2+} to 10^{11} M for Zn^{2+} can all be effectively reversible opposite the respective functions[2]. Thus, in practice, various times of exchange can be compatible with effective reversibility also in enzymes. We note too that a protein fold can be designed so that the metal ion is held irreversibly even though binding is quite weak since the exchange can be dominated by the rigidity of the protein wrapped around the metal ion. The importance of such considerations becomes clear later. All we can do in the absence of full exchange data, which is often the case, is to describe the most likely situation in particular cases. Of course in those circumstances where the enzyme can be extracted without loss of the metal the situation is simplified, since at least we know the precise metal, the use of which we must explain, but now the bound metal ion does not equilibrate with free ion so that we lose the benefit of analysis at equilibrium.

We now turn to a description of one group of enzymes-the hydrolytic (non-redox) enzymes-which provide examples of both different kinds of behaviour, reversible and irreversible. Here we include transferases, hydrolases and lyases in the E.C. classes 2, 3, and 4. By choosing to discuss non-redox enzymes first we avoid the functional value of valence state switches and concentrate upon acid–base properties of divalent cations.

5.3. Selective uptake of metal ions in non-redox active metallo-enzymes: introduction

In this section, as stated, we shall examine particularly the characteristics of the reactions which employ metallo-enzymes in E.C. classes 2, 3 and 4. We stress again that we shall only refer to the divalent metal ions here although occasionally Fe^{3+} and Mn^{3+}, used in acid phosphatases, are mentioned. Divalent ions do not bind very well to proteins at approximately pH 5.0 and hence the use of trivalent ions in these unusual acid conditions. All the enzymes under discussion use water to break covalent bonds in a great variety of substrates including proteins, peptides, phosphate esters and anhydrides, esters, 'ether' bonds of polysaccharides, and amides in

[2] $K(\text{binding}) = k_{\text{off}}/k_{\text{on}}$.

substrates such as urea, or make or break other bonds by acid–base catalysis. The dehydratases, such as carbonic anhydrase, are of a similar character in that they use water to attack substrates. The puzzle concerning all of them is the diversity of metal ions in the enzymes, Tables 15 and 16, when all of them could obviously use just one good Lewis acid metal ion in one simple way-as an attacking acid either on the substrate directly or on a water molecule to produce the attacking base, hydroxide. Why do they use different metal ions? In particular why not use say zinc, which is likely to be the most powerful cation and not open to redox chemistry, in all cases? The answer to this question could lie in localised compartment availability of different ions as described above, or in some particularly useful mechanistic trick, or in an historical evolutionary accident and so on. Again why do some enzymes use two or three metal ions while others do not? Note that we do not include copper here since it may be that the use of this metal could risk damage by redox reactions. We shall also not refer to Ca^{2+} since this ion is used very largely just to bind, even substrates, and is a poor attacking Lewis acid. The selection of calcium in proteins is well understood, based upon thermodynamic binding constants to its ligands, as described above.

A large number of these acid–base metallo-enzymes containing divalent metal ions have been fully sequenced. A fortunate feature is the purity and idiosyncratic nature of the metal content of several of them. Again a sufficient number of them have been isolated without significant loss of metal ions as judged analytically and by maintenance of activity during purification. We may presume that for them exchange is extremely slow and the metallo-enzyme does not equilibrate with free metal ions. Presumably metal ions are bound irreversibly in these cases. Many enzymes of this group have also been examined by X-ray diffraction so as to give a structure which reveals how each metal is bound in vivo, whether it is present as a single atom or in a dimeric or a trimeric centre. Although there remain two areas of uncertainty, the number of water molecules and the exact bond lengths (errors can be hidden and may be as great as ± 0.1 Å, which is not very different from the difference between the radii of zinc (0.6 Å) and manganese (0.75 Å)), the bond angles are usually quite well defined. This situation allows a very good test of whether or not we can understand the selection of metal ion uptake into enzymes based on our knowledge of the structure, on thermodynamic equilibrium binding strengths and/or on the kinetics of binding of metal ions as observed in small molecule model complexes. There is complication in that it is a general supposition that enzymes are 'designed' so as to optimalise functional value. As we have mentioned this implies that catalytic atoms are bound in ways related to their transition state(s) for the cycle of states through which they go during the reactions they bring about. We now use the word 'constrained' (entatic) to describe this state [58]. The question arises as to the compatibility of the two conditions-one based on the best chemical properties for binding the other on the functional value of the metal ions. The best structure for activity may well compromise (reduce) the binding strength.

5.4. Properties of hydrolytic metallo-enzymes

The immediate biochemical information needed concerning these enzymes is therefore:

1. Which metal ion is present?
2. Where in biological space does it act since the metal ion may have its availability dependent on local separation?
3. How many metal ions are present at a given site and why are there more than one in certain cases?

By answering these questions we may then understand why a particular metal is used.

As stated, the metal ions concerned are mainly Mg^{2+}, Mn^{2+} and Zn^{2+}, the major ions being Mg^{2+} and Zn^{2+}. The metal ions Fe^{2+}, Ni^{2+} and Co^{2+} are only infrequently found in these enzymes and one metal ion, Cu^{2+} has never been found in a hydrolytic enzyme. Thus the selection is not based on the ability of the metal ion to assist hydrolysis (see Section 5.3). We consider first the easiest part of the problem posed by this diversity of enzymes-the different uses of magnesium and zinc, the weakest and strongest of the Lewis acids.

5.5. The magnesium enzymes [1,47]

We can immediately separate Mg^{2+} from the consideration of the other metal ions since it is found to catalyse very largely the easiest hydrolyses, e.g. of anhydrides such as ATP, and is selected to bind weakly to either the substrate or its enzyme with a binding constant, $K_M = 10^4$ (Table 17). Virtually all these reactions take place in the cytoplasm. The binding groups of the proteins concerned include perhaps a carboxylate or a phosphate and several weak O-donors such as –OH of serine or threonine. No metal ion has a binding constant above 10^6 to such sites. Given that the concentration of Mg^{2+} is 10^{-3} M and no other metal is present in the cytoplasm at a concentration above about 10^{-6} M only Mg^{2+} will bind to these weak sites (see Figs. 17 and 18). The running of a cell demands a high concentration of ATP (and other nucleotide triphosphates), a high concentration of its enzymes and fast equilibration, so that about 50% of all Mg^{2+} is bound. Note that exchange is fast so that selection of Mg^{2+} everywhere is based upon thermodynamic factors as is the case for Ca^{2+}. However, we must remember that this condition is only applicable while we consider a given part of space, e.g. the

Table 17
Magnesium enzymes [1,2]

1.	Various nucleases mostly in-cells but some are extracellular
2.	Phosphatases both in cells and extracellular
3.	Some DNA and RNA polymerases
4.	Many ATP-ases and synthases
5.	Several enzymes for saccharide isomerism

cytoplasm, in which the free concentrations are set by kinetic pumping, Ca^{2+} 10^{-8} M, Mg^{2+} 10^{-3} M. Outside the cytoplasm enzymes are frequently Ca^{2+} not Mg^{2+} dependent since there the concentrations of both Mg^{2+} and Ca^{2+} are both of the order of 10^{-3} M and the proteins are structured such that Ca^{2+} binds more strongly than Mg^{2+}. There are a few Mg^{2+} enzymes outside cells, e.g. alkaline phosphatases, but there is usually a second metal ion present, e.g. Zn^{2+}. In other cell compartments Ca^{2+} is frequently pumped into vesicles while this is rare for Mg^{2+} (compare Fe^{2+}). We conclude that Mg^{2+} is used as the catalytic ion for reactions where hydrolysis is relatively facile and there is plenty of magnesium available, that is mainly in the cytoplasm. For this reason magnesium assists in the coordination of control over many substrate pathways of ATP and PAPS (see Fig. 24). The exchange of magnesium extends to the major phosphate controls of DNA expression and undoubtedly [Mg^{2+}] in cells is constitutively regulated (see Section 1.3 and Fig. 24).

Now there are also several examples of Mg^{2+} hydrolytic enzymes where there are either one or two metal ions at the active site, e.g. in nucleases. The ribose-phosphate bonds are less easily broken than those of pyrophosphate and these substrates are also less easily bound by magnesium. The use of magnesium here can be put down to either effective catalysis and/or to a functional control use in linking different reaction paths or in links to DNA expression of proteins since Mg^{2+} exchanges from virtually all its bound sites (see chlorophyll, below, for an exception).

As mentioned earlier, for an ion with many linked equilibrating functions, the binding constants for all these sites, single or multi-nuclear, must then be closely the same as that to ATP i.e. $\log K = 3.5–4.0$. N.B. equal binding strengths of all sites is necessary if the sites are to be equally occupied at equilibrium.

5.6. Chlorophyll proteins

We turn now to the first clear-cut case of non-equilibrium binding, that of Mg^{2+} in chlorophyll. The binding of Mg^{2+} to five N-donors, four of a chlorin ring and one of an imidazole side chain of a protein, is not possible at equilibrium in any cell in the presence of even 10^{-11} M zinc or 10^{-8} M Fe^{2+} ions. In fact there is a known irreversible insertion pathway using a chelatase in a chloroplast compartment. The reaction requires recognition (reversible or not) of Mg^{2+} or a Mg^{2+} complex to the exclusion of other ions due to binding selectivity as far as possible, followed by a selective physical transfer of Mg^{2+} into chlorin. The overall transfer has to be specific due to a combination of the binding constant and shape selectivity. For example ATP binds Mg^{2+} alone in cells. Thus Mg · ATP could be selectively recognised, and if necessary using the energy of ATP, the Mg^{2+} could be forced to bind to chlorin. A further selection can take place as chlorophyll, Mg · chlorin, is recognised by a binding protein and then inserted, by a mechanical transfer, into a membrane. The process of chlorophyll protein formation is irreversible (see Section 7.1).

5.7. The spatial location of metal enzymes other than those of magnesium [1–3]

It is obvious that bio-organic chemical systems, consider DNA, cannot tolerate even quite low concentrations (10^{-6} M) of metal ions such as Mn^{2+}, Co^{2+}, Zn^{2+} and even Ca^{2+} (i.e. all other than Mg^{2+}) as they may damage this genetic material. As binding and Lewis acid catalytic efficiency increase so the metal ion, even Zn^{2+}, becomes more dangerous. These ions are, however, essential for organisms in various enzymes, as listed in Tables 15 and 16 and the ions must therefore be transported, at equilibrium or not, to the regions where the enzyme is synthesised or, for those cases in which the enzyme is transferred, to where the activity of the enzymes must occur. At the same time the excess of the metal ion must be removed from the localities in which it could do most damage, e.g. around DNA. One possible solution is that some metal ions and their proteins are positioned in cells by transfer, that is by pumps (exchangers), carriers and/or channels into compartments away from the cytoplasm. Analysis shows the disposition of metal ions in cellular systems of all kinds Table 14. Notice that even small excess of any metal over its proteins is prevented by uptake inhibition or export into extracellular space or vesicles, the latter in eukaryotes only of course. The vesicle systems into which the proteins and metal ions can go include storage sites such as the plant vacuole, vesicles for discharge, such as the adrenal or signal peptide vesicles, the endoplasmic reticulum, the Golgi apparatus, synaptosomes, vesicles for uptake, such as lysozomes, and those for protection such as peroxyzomes. Again cellular organelles pick up particular metal elements and proteins. The metal and protein content of each of these local systems is remarkably selective. The management of the distribution ensures that each metal is finally bound specifically and separately from all others to a given protein, X, possibly to generate an enzyme, in an appropriate place. This selection may not have been made at equilibrium even in an early transfer step, but may have occurred using the shape of a metal ion complex to bias transfer, often using energy. The transfer is therefore based on kinetically controlled passage through membranes or even through the cytoplasm, Fig. 3, using a specific carrier, Y, or a pump, Z, to decide the region of the cell in which the metal ion is to be located before final insertion in X. Y (or Z) then also has two functions: it selects a metal ion by binding constant and by shape and it then constrains each metal to a 'specific' transfer route or pathway. In earlier sections we have treated transfer at equilibrium but we must revise this view with respect to many enzymes.

We turn to the description of particular metal ions, other than Mg^{2+}, and their enzymes.

5.8. Zinc extracellular enzymes [52,56]

The vast majority of hydrolytic enzymes contain non-exchanging zinc and are extracellular. Its value based upon its known inorganic chemistry has been stressed many times [59]. Enhancing its value, in what is now to be called a constrained (entatic) state within the protein matrix, is dependent upon activating water, as hydroxide, or the substrate by the acidic character of this cation which has been

Table 18
Zinc enzymes [1,2,52–56,66][a]

Class	Coordination	Spatial distribution
Carboxypeptidases	H E H	Vesicles, extracellular
Aminopeptidases	H E H	Vesicular, extracellular
Alkyl phosphatase	D S D H	Extracellular
Nucleases	W H D H H D	Extracellular
Collagenase	H H X, [(C) or (H_2O)]	Extracellular
Phospholipase C	H H E (H_2O)	
	H W D (H_2O)	?
	H H D D	
β-Lactamases	H H D H H	Extracellular
Glyoxalases II		
Aryl sulfatases		
Phosphohydrolases	H H D D	Extracellular
Leucine amino peptidase	K D D D E	Intracellular (?)
Calcineurin	D N H H (H_2O)	Intracellular
Aspartate trans-carbamylase	C C C C	Intracellular
RNA/DNA Polymerase	C C H H	Intracellular

[a] See Ref. [57] and references therein.

increased especially by protein structural constraints. Certain constraints are due to the rigid β-sheet structures of the proteins, which also prevent or restrict exchange (half-life of many days). All of these features appear to be well understood and apply to a remarkable range of enzymes in many different zinc hydrolytic classes, Table 18. The zinc then may be combined irreversibly with proteins on folding in vesicles (enzymes) and contained in vesicles ready for export. Finally exported zinc enzymes with irreversibly bound zinc may be placed on the external face of the cytoplasmic membrane where they act as agents for production of peptide hormones, e.g. angiotension converting enzyme, or destruction of signal peptides. It would appear in fact that zinc is the general and virtually the only powerful divalent catalytic metal ion used for such reactions in *extracellular space*, and we know of no case where the divalent ions of Fe, Co or Ni are so used. The favouring of zinc probably relates not so much to its great catalytic power, which is often matched on substitution by one or two of the other cations, but by the absence of redox activity and by its greater binding strength to the proteins, which helps to prevent exchange. Generally Mn^{2+}, Fe^{2+} and Co^{2+} bind more weakly and exchange more readily. In some of the enzymes there is more than one zinc ion and now and then zinc is used together with Mg^{2+}, where, as stated earlier, Mg^{2+} acts to assist binding more than it helps in direct catalysis. We shall consider again this use of zinc after we have looked at evolution of the use of other elements.

In some vesicles the zinc may even be very loosely bound at concentration levels approaching 10^{-6} M for example in insulin, mossy fibre synaptosomes, or reproductive tract vesicles, but here zinc is needed as free, released zinc in transmission.

We turn now to the few zinc enzymes of the cytoplasm. We treat zinc redox enzymes which employ co-enzymes such as NAD together with acid–base enzymes as zinc itself cannot be involved in redox steps).

5.9. Zinc in the cytoplasm

There is a small number of zinc enzymes in the cytoplasm but they appear to be curiously unrelated to those in the extracellular space. One outstanding feature is the association of zinc with DNA/RNA synthesis and degradation which extends from viruses to higher animals and appears to have been present very early in evolution. These enzymes, nucleases in effect, utilise Zn^{2+} binding through N/S centres, Table 19. They are therefore related to cytoplasmic dehydrogenases and aspartate transcarbamylase. Other four coordinate sites of zinc enzymes of the cytoplasm include superoxide dismutase (4N-donors) and carbonic anhydrase (3N-donors and one water). There is little reason to suppose that these enzymes equilibrate with free zinc ions since the proteins have strong β-sheet folds even in the apo-enzymes, and the rates of exchange are very slow. A different class of zinc proteins, not now enzymes, are the zinc finger DNA-binding proteins, which are transcription factors for hormones and we consider for free zinc, and metallo-thioneins. Although they are bound strongly to N/S donors they exchange quite readily and their apoproteins are random coils of low fold stability, Table 20. Thus with two remarkable exceptions, those of calcineurin and leucine aminopeptidase (see below), we find that all the zinc of proteins in the cytoplasm is bound similarly (N/S binding) and quite differently from the zinc for use in exported proteins (N/O binding). N.B. It may be that zinc does not bind powerfully enough to RS^- groups to prevent oxidation to –S–S– cross-links when outside the reducing atmosphere of the cytoplasm.

Now the two exceptional examples of zinc incorporation cross the boundary of coordination of what we expect for Mg^{2+} (perhaps Mn^{2+}), all O-donors, and those

Table 19
Zinc enzymes in the cytoplasm [1,2]

Enzyme (protein)	Coordination
DNA-related and RNA-related phosphatases and synthetases	N/S, possibly four-coordinate structural function
Aspartate transcarbamylase	N/S four-coordinate. Structural
Alcohol dehydrogenase (many NAD-dehydrogenases)	N/S structural and active sites
Superoxide dismutase	4N structural
Carbonic anhydrase	3N active site
Leucine amine-peptidase	O four-coordinate
Calcineurin (phosphatase)	D, N, 2H(H_2O) active site
5-aminolevulinic dehydratase	3S functional
P-53	N/3S tetrahedral structure
Protein kinase C	4S structural

Table 20
Rates of dissociation of zinc [2]

Protein	Zinc half-life of exchange
Angiotension-converting enzyme (extracellular)	36 s
Carboxypeptidase (extracellular)	29 min
Astacin (extracellilar)	39 days
Carbonic anhydrase (?)	3 years
Metallo-thionein (Cyt.)	2 s (3-cluster) 10^{-3} s (4-cluster)
Zinc fingers (Cyt.)	~1 s

expected for Fe^{2+}, Co^{2+} and Ni^{2+}, N-plus O-donors. How is it possible for such a relatively slight diversity of sites to bind selectively a diversity of metal ions in the same cytoplasmic compartments? Some very similar sites of oxido-reduction enzymes containing Fe^{2+}, Co^{2+} and Ni^{2+} are also very specific in their protein associations (see Section 6) and clearly are protected from Zn^{2+}. We note that Mn^{2+} and Fe^{2+} are weakly binding ions which can exchange from their sites. Thus these sites in particular need protection from free Co^{2+}, Ni^{2+}, and Zn^{2+}. We therefore go to the description to all these elements in their hydrolytic functions in the cytoplasm. Before turning to them we note that the ways of distributing zinc to the different zinc functions is only just beginning to be clarified, Fig. 25 and free $[Zn^{2+}]$ is uncertain. Since there is only one Ni^{2+} hydrolytic enzyme, urease, found in a vast range of species we consider this example first.

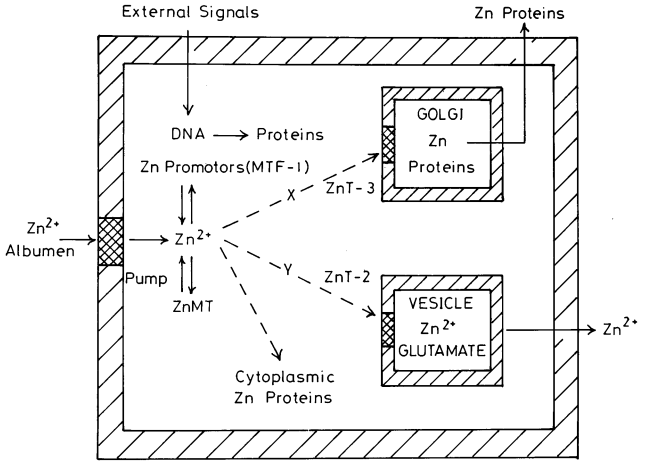


Fig. 25. The selective uptake of zinc and its binding to proteins. In the cytoplasm there are many slowly reversible zinc proteins, binding constants around 10^{11} M^{-1} but in some vesicles and in extracellular fluids zinc is irreversibly bound. Even free zinc at 10^{-4} M in other vesicles does not exchange with the cytoplasm though it equilibrates in these signalling systems. X and Y are carriers.

5.10. Nickel in urease [51]

There is little doubt today that nickel is handled in the cytoplasm of cells so that it never equilibrates with free nickel ions. There appears to be a basic hand-on of nickel from nickel pumps, which must exchange at some stage with the extracellular environment, to carriers and then to the only nickel hydrolytic enzyme, urease, present in some bacteria and some (not all) cells of advanced eukaryotes, e.g. soy bean. One at least of the carriers is a transcription factor informing the DNA of demands for urease synthesis, presumably coupled to transcription factors linked to urea itself or its product ammonia but not to free nickel.

Given the information from model chemistry, Section 3, and the known fact that plants which accumulate nickel do so by carrying it to cells as a histidine complex in extracellular fluids [40] it may well be that Ni^{2+} pumps are based on N-(histidine) donors and not on N/S donors, contrast zinc. We also know that excess Ni^{2+} is largely removed into vacuoles and is passed into them often together with Co^{2+} but with little if any Mg^{2+}, Fe^{2+}, Zn^{2+} or Cu^{2+} (or Cu^{+}). In the vacuole Ni^{2+} can be as high as $\sim 10^{-5}$ M and is only bound weakly often by dicarboxylic anions if at all. Thus, Ni^{2+} is handled irreversibly inside the cytoplasm of plant cells. (Note that it is not used in animal cells). By handling nickel in non-exchanging complexes the problem of equilibrium exchange with zinc for example which must occur in the earliest steps of uptake, e.g. in formation of extracellular histidine complexes, is avoided (see Fig. 26).

The difference in shape of the nickel and zinc complexes, in proteins or not, allows them to be transferred differently. The later handling of nickel in bacteria and plants, non-exchanging, then parallels that of sulfur, not that of exchanging magnesium and calcium. Even in *E. coli* there are five gene products for the handling of nickel [27].

Nickel is also used in a special complex, F-430, in archae bacteria.It may well be that the handling of nickel and cobalt is similar in many respects in these organisms. It is certainly the case that Ni and Co are both put irreversibly into ring chelates, F-430 and vitamin B_{12}, but the reactions of these centres are not hydrolytic.

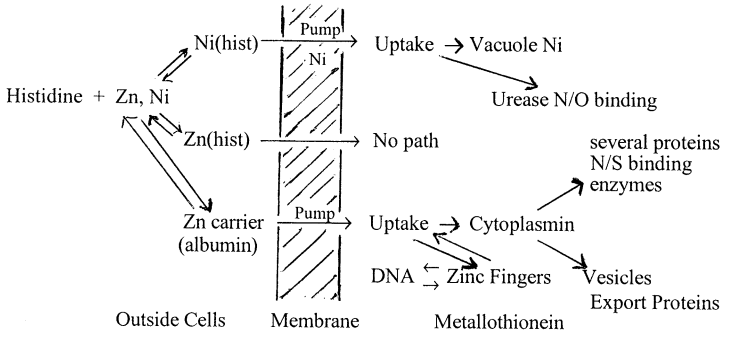


Fig. 26. The contrasting pathways of selective uptake of zinc and nickel in plant cells.

Table 21
Cobalt proteins[a] (excluding B_{12}-dependent) [38,50]

Proteins	Function and comment
Methionine aminopeptidase	Removal of terminal methionine Co(II) dimer (bacteria to animals)
Prolidase	Hydrolysis adjacent to proline Co(II) dimer
Glucose isomerase	Unusual: in actinomycetes. Other M(II) cations are active
Methylmalonyl-CoA Carboxytransferase	Bacteria (Co(II) + Zn(II))
Aldehyde decarbonylase	Algae
Lysine 2,3 amino-mutase	(Bacteria); Zn(II), Fe_n/S_n also
Bromoperoxidase	Bromination (bacteria); V or Fe in other organisms (e.g. algae)
Nitrile hydratase	–CN bond hydration (bacteria) Co(III). Sometimes Fe(III) enzyme
Cobalt transporter COT 1, GRR1, HoxN	Transfer of Co(II) (bacteria and yeast). N.B. Note activity with Zn(II) and Ni(II)
Cobalt porphyrin protein	Electron-transfer? (bacteria)

[a] N.B. Note the absence of plant proteins and the rare appearance in animals. All proteins are in the cytoplasm.

5.11. Cobalt hydrolases [37,60]

Some cobalt enzymes are listed in Table 21. The curiosity concerning cobalt is its link to methionine peptidases in the cytoplasm. Now, this activity is not recorded in higher plants but in most if not all other species. It is required in them for the removal of the N-terminal amino acid, methionine, which is an initiator and an initial amino acid of protein synthesis. Such a fundamental link between cell metabolism and cobalt is related to other functions of both methionine and cobalt in other reactions. Thus coenzyme B_{12} (cobalt) and S-adenosyl methionine are both used in methyl transfer reactions. Moreover the synthesis of methionine itself is vitamin B_{12} dependent. Hence there is a tight loop of cobalt/methionine dependencies in some lower species and in higher animals. Clearly the incorporation of cobalt into vitamin B_{12} requires an insertion reaction to avoid Zn^{2+} which binds better than Co^{2+}. Therefore it is entirely possible as stated that cobalt like nickel is handled in the cytoplasm quite independently from free ionic equilibria. In fact the synthesis of cobalt B_{12} requires some 15 Cob proteins. Note that the problems for cells of all these insertions into F-430 (Ni), vitamin B_{12} (Co), heme (Fe) and chlorophyll (Mg) were solved extremely early in evolution (see Section 8.4).

Returning to the enzymes which contain cobalt ions bound simply to protein amino acids, the cobalt distribution has been studied in several bacteria. The outline uptake system is given in Fig. 23 and in essence is similar to the general uptake scheme for metal ions which are essential, Fig. 3. There is at least one pump for import into the cytoplasm and one for export either to extracellular and/or to vesicular compartments. In the cytoplasm there is at least one carrier/transcription factor which controls several genes for the expression of all the cobalt dependent systems. All of these cobalt proteins could communicate in principle through

selective cobalt binding at equilibrium or they could cooperate through kinetically controlled transfer between themselves — an insertion mechanism. The most likely process is that of non-equilibrium transfer since it is difficult to see how competition for the N/O donor sites of enzymes such as those of methionine peptidase could avoid competition (at equilibrium) from zinc and iron. In higher organisms there may be carriers of cobalt in the extracellular fluids which interact with free cobalt and uptake cell pumps much as is described above for internal carriers.

5.12. Iron hydrolases [49]

Iron, Fe^{2+}, in the cytoplasm is overwhelmingly used in oxido-reduction reactions, Section 6. Its handling, largely reversible, in *E. coli* is shown in Fig. 27. There is one example where the function is in an acid–base enzyme and a transcription related factor-calcineurin in higher organisms. Here Fe^{2+} and Zn^{2+} work together to form an extremely important calcium-regulated phosphatase. The function of the Fe^{2+}, which can exchange, is quite possibly as both a binding group, zinc being the attacking group, and as a control opposite the redox status of the cell. The ferric (Fe^{3+}) enzyme is inactive. The Fe^{2+} binding is quite weak (and the exchange is quite fast) indicating that this is too weak a site to hold zinc at its level of 10^{-11} M-the equilibrium level, while that of Fe^{2+} is $>10^{-8}$ M. It could hold Mn^{2+}, but not Mg^{2+}, and it maybe that in many, especially plant calcineurin phosphatases, Mn^{2+} is the active metal ion. Mn^{2+} levels are much higher in plants and Mn^{2+} calcineurin is active.

If this description proves to be correct and since Fe^{2+} is generally in fast exchange then it, like all its other sites, needs protection from free zinc. The protection is provided by the generally low binding constant of the site and

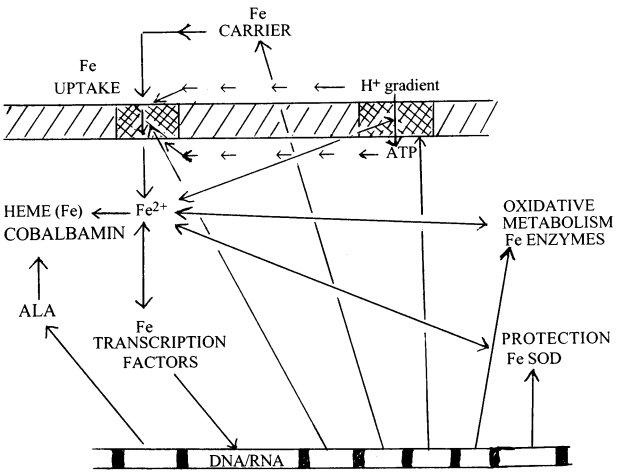


Fig. 27. The selective uptake of iron, Fe^{2+}, is based on reversible binding constants at around 10^{7} M^{-1} and like Mg^{2+} it has a large role in the cytoplasm but a much smaller role elsewhere, see Section 6.2.

relatively small difference in binding constant between Zn^{2+} and Fe^{2+} generated by a site of octahedral coordination and open sidedness (see Fig. 4) comparing EGTA larger site and EDTA smaller site. Thus the lowering and limitation of Zn^{2+} to 10^{-11} M by N/S sites of metallo-thionein and zinc fingers and the irreversible coordination described above prevents it from competing with weaker sites and Fe^{2+} at 10^{-8} M.[3]

5.13. The manganese enzymes [48]

The manganese enzymes are listed in Table 22. As for cobalt and nickel the use of manganese in peptidases in the cytoplasm appears to be arbitrary. Manganese is used in both proline peptidases and in arginase, but why such substrates require this enzyme is unknown. The active sites of its enzymes are particularly well adapted to the selection of Mn^{2+} rather than the other ions of the first transition metal series since they are mainly based upon O-donors and give roughly octahedral sites, but compare zinc in leucine amino-peptidase.

A striking feature is that two hydrolytic enzymes one using nickel (urease) and the other using manganese (arginase) are needed in the urea cycle, Fig. 28. The question then arises as to why zinc or magnesium is not used? It is very likely that Mg^{2+} is too weak a catalytic acid to be effective. The avoidance of zinc is more probably to be an historical accident due to the fact that zinc was less available earlier in evolution. As stated zinc in the cytoplasm became retained at levels $< 10^{-11}$ M by N/S sites which are not used in hydrolyses. Such sites may not be useful in simple acid/base catalysis due to the reduced acidity of four-coordinate Zn^{2+}.

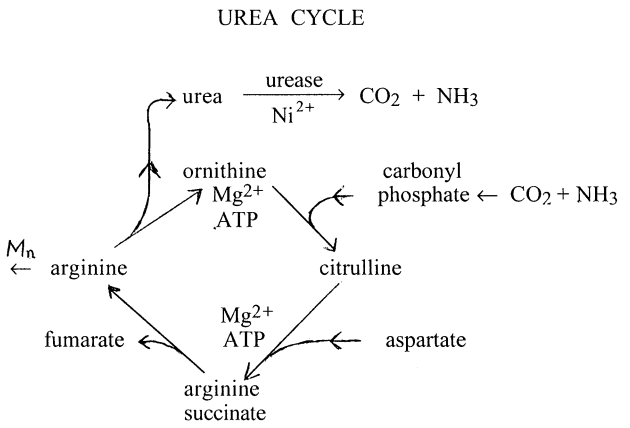


Fig. 28. The urea cycle of plants and lower organisms not animals. Note the roles of Ni^{2+} and Mn^{2+}.

[3] A newly discovered Fe^{2+} hydrolytic enzyme in bacteria, formyl hydrolase, acts to remove the formyl group of the initiator of protein synthesis in bacteria (and organelles) which is formylmethionine. It appears as if the cytoplasm prefers Fe, Co and Mn over Zn in hydrolytic enzymes. Is this an historical feature?

Table 22
Mn enzymes and proteins

Protein	Valence	Binding site	Function
Concanavalin A	II	Glu, Asp, Asp, His H_2O, H_2O	Saccharide binding
Proline dipeptidase E.C.3.4.11.5	II	O and N ligands proposed	Hydrolyses C-terminal proline dipeptides (ubiquitous)
Galactosyl Transferase E.C.2.4.1.38	II	Mn specific in vivo	Transfer of galacto-group (Golgi)
Arginase E.C.3.5.3.1	II	4 Mn atoms per enzyme molecule	Hydrolyses guanidino-acetate to ornithine and urea (mitochondria)
Glycocyaminase E.C.3.5.3.2	II	S–H group near active site	Hydrolyses guanidino-acetate to glycine and urea (mitochondria)
Phosphoenolpyruvate carboxykinase E.C.4.1.1.32	II	Binds through O. N.B. Fe competes not Mg.	C-C carboxylyase (mitochondria)
Pyruvate carboxylase E.C.6.4.1.1	II	4 Mn atoms per enzyme molecule	C-C carboxylating synthetase (mitochondria)
Isopropylmalate synthase E.C.4.1.3.12	II	S–H group near site	C-C oxo-acid lyase (mitochondria)
Acid phosphatase E.C.3.1.3.2	III	His residues essential	Hydrolyses *ortho*-phosphoric monoester (prokaryote)
Mn-ribonucleotide reductase	II		Reductive elimination of 2-hydroxyl from ribonucleotide (prokaryote)
Mn-dioxygenase	II	Low symmetry Mn(II) centre	Catalyses extradiol cleavage of catechols (prokaryote)
Mn-catalase E.C.1.11.1.6	III		Redox enzyme; acts on H_2O_2 (prokaryote)
Mn-peroxidase E.C.1.11.1.7	II/III		Redox enzyme as above, degrades lignin (extracellular)
Superoxide dismutase E.C.1.15.1.1	III	His, His, Asp, His, H_2O	Redox enzyme; acts on superoxide radical (mitochondria, chloroplasts, prokaryotes)
(Mn-pyrophosphate	II/III		Acts in place of SOD in lactobacillus
Photosystem	II/IV	Assumed typical N/O donor site	O_2 release in plants (chloroplasts)

A further feature of Mn^{2+} in enzymes is that they are found in the Golgi and in organelles. Thus Mn^{2+} is largely pumped out of the cytoplasm to levels around $< 10^{-6}$ M while $[Mn^{2+}]$ in vesicles may be much higher, e.g. in plant vesicles. The use of Mn^{2+} in galactosyl transferase in the Golgi is not easily explained since zinc should be much more effective unless this too is a remnant of evolution. Neither

Mg^{2+}, nor Zn^{2+} nor Fe^{2+} are pumped into the Golgi and the only possible competition is from Ca^{2+} which is not known to bind N-donors.

We can find no other simple explanation for the various uses of Mn^{2+}, Fe^{2+}, Co^{2+} and even Ni^{2+} in the cytoplasm in acid–base enzymes and the little use of Zn^{2+} whilst outside cells Zn^{2+} is used extensively. We examine this situation again in Section 8, where the availability of elements is considered in an historical context.

5.14. Metal clusters and multi-metal ions in proteins and enzymes

Apart from simple clusters such as Fe_2O, Fe_4S_4 and the clusters in metallothioneins and some zinc fingers of metal ions with RS^- groups there are a large range of enzymes which have several metal ions Table 16. We have explained the possible enzymic functional significance of these structures in Section 5.2. Now there is also the consideration of their stability and their exchange rates. As we have stated repeatedly it is desirable to have relative fast exchange for sites which are involved in messages internal to a cell or in buffering. However the binding in the sites which exchange has to be of approximately the same strength as that of the buffer of the particular metal ion. The alternative is to lock the metal ion in a non-exchanging site, e.g. iron in heme. In those cases where there is more than one metal ion it may well be that the binding of the second or third ion changes the status of the first or even both ions from an exchanging to a non-exchanging site. The kinetics then allow that only when the two required metal ions bind is the system irreversible. The use of the specific shapes of preferred coordination spheres and the most stable fold of the protein will allow different properties of combinations of metal ions to generate quite different functions. For example it can be that the curious coordination of two nickel ions in urease, of three zinc ions in phospholipase C, and of iron and zinc in calcineurin derives from the cooperative combination of the metal ions and the proteins in a rather complicated manner.

In conclusion we consider that the distribution of metal ions in acid/base reactions in organisms can only be explained on the basis of the division between thermodynamic equilibration of some sites and kinetic limitation of exchange at others, according to the following pattern.

Zn → largely irreversible tight binding, many cases for export

Ni, Co → largely irreversible not quite so tight binding, cytoplasmic

Fe, Mn, Mg → largely reversible weaker binding in different compartments but not for export. Fe and Mg largely cytoplasmic

These generalisations apply to the cytoplasm in particular. We shall consider later that this division is an historical accident.

Table 23
The cytoplasmic protein-phosphate phosphorylases

Phosphorylase	Metal ions	Donor atoms (C.N.)
P P 1	Fe^{2+}	H, H, D, D, $2H_2$ (6)
(P P 1 C)	Mn^{2+}	H, D, D, $2H_2O$ (5)
P P 2 B	Fe^{2+}	H, D, D, $3H_2O$ (6)
(P P 2 A C)	Zn^{2+}	H, H, N, D, H_2O (5)
P P 2 C[a]	Mn^{2+}? Mg^{2+}	D, G, (H_2O) (6)
	Mn^{2+}? Mg^{2+}	D, D, D, $(H_2O)_{2/3}$ (6)
P P 1 B	No metal	
(P T P)	Ions	

[a] Present in archae bacteria.

5.15. The cytoplasmic protein-phosphate phosphatases: genetic controls

There is a peculiar feature of the cytoplasmic protein-phosphate phosphatases. They are often metallo-enzymes but the metal content of the different functional units contain some five different metals, Mg, Mn, Fe, Zn and Ca [3]. Their coordination chemistry is shown in Table 23 and agrees with expectation from model complexes. The metal ions Mg^{2+}, Mn^{2+}, Fe^{2+} and Ca^{2+} are known to be in rapid exchange while Zn^{2+} is more firmly but unusually held and is probably in exchange at slow rates. The enzymes act on different protein phosphates and all control either metabolism or expression through transcription factors. For example the Fe, Zn enzyme controls part of the immune system response while the Mg controls the expression of the calcium ATP-ase pump. The inclusion of a calmodulin Ca^{2+} binding site in the Fe/Zn phosphatase, the presence of Zn^{2+} in it and its function all indicate that it is not a primitive enzyme. There are no calmodulins in prokaryotes and it appears that zinc is a relative new-comer in the cytoplasm. The other phosphatases appear to be primitive but only one has been found in archae bacteria.

The network of communication, phosphorylation/dephosphorylation, in the cell is then very extensive. In fact about one third of mammalian proteins are phosphorylated and the human genome may well have 2000 protein kinases and 500 protein phosphatases and this central role of phosphate transfer in signalling as well as energy transduction is apparent from the earliest organisms. It is of very considerable interest then that the phosphorylation/dephosphorylation control system over protein activities is linked to four different dissociable metal ions since this neatly connects it to the control systems of these four metal ions. Note that the internal phosphatases all use divalent ions while the similarly structured external acid (purple) phosphatases use Fe^{3+} (or Mn^{3+}) and have a metal-tyrosine bond. (Some vanadium dependent haloperoxidase have a similar fold).

6. Redox enzymes: distribution of metals

6.1. Introduction

The chemistry of redox enzymes has been well described recently [62] but the description largely ignores the problem of their selective formation.

There would appear to be relatively few cases in which the particular metal ion in a redox enzyme, classed in E.C. Group 1 (oxido-reductases), is selected by thermodynamic equilibrium competition between metal ions for a given site since quite often the metal does not exchange from any of these sites [61,62]. We have already mentioned the iron heme, cobalt corrin, magnesium chlorin and nickel F-430 ring systems as examples of metal ion insertion in the co-factors of a variety of redox enzymes, from which there is virtually no exchange. Other examples are the formation of the two molybdenum co-enzymes from which molybdenum does not exchange, most of the copper enzymes where binding is too tight to permit exchange, and nickel hydrogenases from which nickel is not known to exchange and formation is due to carrier/insertion systems. The overall implication must be that free metal ions except Fe and Mn do not control the synthesis (at the DNA level) of these proteins and that this control is exerted by co-enzymes for example of iron, such as heme, or those of molybdenum. In other cases it may be the concentration of the inserting protein which controls the formation of the enzyme.

Whereas the replacement in an acid/base catalyst of one divalent metal ion by another taken from the Irving–Williams series gives only a modest change in acidity function, whence activity can be expected and is found for a series of different metal ions in enzymes in vitro, switching metal ions at oxidation-reduction centres is unlikely to be successful. It would appear that the identity of a metal ion for a given redox enzyme should be the same in all organisms. (An apparent exception is the exchange of Zn for other metals in dehydrogenases. We shall not refer to these zinc enzymes again in this section since the function of the zinc is as a Lewis acid and the enzymes are already included in Section 5.8). It is then surprising to find cases of different metals acting in the same redox reactions listed in Table 24 [63], though we do not know whether there is exchange in most of these cases. It is known, however, that the concentrations of Fe and Mn superoxide dismutase are controlled by the concentrations of free metal ions both directly and through the transcription factor for metal ion uptake, Fur, which does exchange metal ions readily.

Table 24
Redox proteins with low metal ion selectivity [63]

Superoxide dismutase	Mn, Fe, Cu/Zn
Ribonucleotide reductase	Co(B_{12}), Fe, Mn
Catalases	Fe, Mn
Peroxidases	Fe, Mn, Se
Aldehyde reductases	Mo, W
Nitrogenase	(Fe)Mo, V

Two notable features of Table 24 are the use of Mn and Fe on the one hand and Mo and W on the other in the same reactions [62]. It so happens that on the oxidising side of $+0.0$ V it is relatively easy to obtain the same one electron potentials for Fe and Mn and the same two electron potentials for Mo and W so that the use of different metal ions here may be of little consequence. More surprising are the use of Cu in place of Fe/Mn in superoxide dismutases and of Se (or V) in peroxidases rather than heme iron. We believe that the first is an example of an evolutionary gain of protection from exchangeable Mn or Fe for non-exchanging Cu, while the second is the gain due to the ability of Se to undergo reaction with classes of secondary and tertiary peroxides different from those open to attack by heme enzymes.

A second problem, which is specific to those elements which can undergo oxidation and/or reduction to give different oxidation states, is that the redox states of these elements in different enzymes are found not to equilibrate with one another. We met this problem in the case of hydrogen when we described which conditions of bound hydrogen were in equilibrium with H^+, Section 4.6. We know for example that the potential of NAD/NADH and flavin couples do not equilibrate with the potential of quinones Q/QH_2. Now these redox couples are at the ends of the energy capture chain of anaerobic organisms. They communicate through various other redox couples in membranes including those of iron. The NADH couple gives electrons to Fe/S proteins in particle I with redox potentials around -0.4 V which then pass electrons to particle III cytochromes and Rieske proteins at the quinone potentials around $+0.1$ V. The several different iron proteins in the series cannot be in redox balance. Some are even on opposite sides of membranes. We shall consider that only those iron proteins which exchange free iron easily in the cytoplasm are of necessity in redox equilibrium with one another and with the cytoplasmic redox buffer, glutathione. The implication again is, much as it was for hydrogen, that the total iron in a cell is not related to the free iron and that only some proportion of it is in equilibrium with the cytoplasmic redox potential. Clearly the iron in these energy capture proteins is not in exchange equilibrium.

This leaves us with the cases of enzymes where we know that the metal ion is readily lost from the redox protein during isolation, i.e. it exchanges readily and hence must be in equilibrium with the free metal ion in a cell [64]. We consider these cases next. Before doing so we observe that, just as it was the case amongst hydrolases, transferases and lyases, the use of the metal ions of nickel and cobalt appears to be disappearing in aerobic higher organisms [2,3] so it is true that the only nickel oxido-reductases, four of them, are present in prokaryotes and archae bacteria, (but not in eukaryotes) and appear there to be irreversibly formed. The redox use of cobalt is almost confined to irreversibly formed vitamin B_{12} derivatives, which do not appear in plants, and there would appear to be very few manganese oxido-reductases in higher animals, perhaps none in the cytoplasm. Manganese is, of course, more dominant in plants especially through O_2-producing centres. Here the manganese may exchange slowly in the chloroplasts. Apart from the use of V and Mo in special co-enzyme centres, where the co-enzyme not the

Table 25
Coordination of Fe^{2+} in cytoplasmic enzymes [1,2,61,62]

Enzymes	Donor groups	Coordination[a]
Hemerythrin	3His, Glu, Asp	6-coordinate
	2His, Glu, Asp	5-coordinate (?)
Ribonucleotide Reductase	His, 2Glu, Asp	6-coordinate (?)
	3Glu, Asp	6-coordinate (?)
Methane Monoxygenase	His 3Glu (2 atoms)	6-coordinate (?)
Fatty Acid Desaturase	His 4Glu (2 atoms)	6-coordinate (?)
Extradiol Dioxygenase	2His, Glu	6-coordinate (?)
NO Reductase (Membrane)	3His	6-coordinate (?)
Reaction centre (membrane)		6-coordinate (?)
Calcineurin	His, Glu, Asp	6-coordinate
Penicillin synthetase (IPNS) [67]	2His, Asp, Asn($2H_2O$)	6-coordinate
2-Oxyglutamate using oxidases	2 His Glu (Asp)	6-coordinate
Regulatory proteins (Fur)	His(His) Glu?	6-coordinate?
Diphtheria toxin repressor	His Glu His	6-coordinate

[a] 6-coordinate structures use some water molecules and the groups are usually in roughly octahedral distribution.

metal may exchange, this leaves only the problems of the common distribution of Fe^{2+} (Fe^{3+}) generally and perhaps some Mn^{2+} (Mn^{3+}) in equilibrium balance in redox active proteins. It is noteworthy that the vast majority of Fe^{2+} redox enzymes are in the cytoplasm while those of Cu, Mn and, in higher organisms, Mo, are in vesicles, organelles or in external fluids. Exchange is not valuable outside cells but is extremely useful inside cells since there it can control DNA activity, apart from acting as a homeostatic device amongst metallo-proteins and enzymes. We have seen that free iron does equilibrate with transcription factors (Section 4.4).

6.2. Fe^{2+} sites in proteins

Iron is the main example of an exchangeable metal ion in a redox enzyme in the cytoplasm. Such sites are then at risk from preferential binding by Zn^{2+} in particular [2,65]. Table 25 shows however that many sites occupied by Fe^{2+} ions differ from those binding Zn^{2+} in four respects:

1. Fe^{2+} is usually in an octahedral site (Table 25), while Zn^{2+} is in 4- or 5-coordinate sites even though the binding ligands are two histidines and one carboxylate in both cases. The bond angles and bond lengths may then be different [67] and the binding to these centres is relatively weak.
2. Fe^{2+} is never bound directly to a combination of N- and RS^--donors in single metal ion sites, contrast the major cytoplasmic zinc sites in the cytoplasm, Table 12. N.B. Some Fe^{3+} sites are due to $4RS^-$ groups, see below). Here zinc exchanges slowly.
3. Fe^{2+} binds reversibly to Fe_3S_4 units in vivo even in some transcription factors, while Zn^{2+} does not to do so [65].

4. While some binding sites for Fe^{2+} have the same combination of N- and O-donor groups [68], as do many sites for Zn^{2+}, overwhelmingly the Zn^{2+} sites are outside cells while those of Fe^{2+} are in the cytoplasm, Section 5.8.

Generally all the Fe^{2+} sites are in the cytoplasm not in vesicles or extracellular fluids and readily exchange iron and lose it on isolation.

Now the binding of Fe^{2+} to most organic molecule (protein) centres is relatively weak $K < 10^8$. Using model data the binding of Zn^{2+} to such a site, where the size of the binding site is optimal for ions larger than Zn^{2+}, e.g. EGTA (see Fig. 15) is expected to be only some 10^2 stronger in octahedral symmetry. For the Fe^{2+} enzyme to be correctly and safely formed in solution in the cell therefore demands a concentration of free Fe^{2+} about 10^3 times greater than that of free Zn^{2+}, which we believe to be close to or less than the difference in cellular levels occurring naturally, Note: by making the cavity smaller and tetrahedral and therefore stronger for Zn^{2+}, this ion would be preferred (see Section 3). Thus it is the limitation to larger cavity size due to the protein fold which allows Fe^{2+} binding in the presence of Zn^{2+} to certain sites while allowing the reverse preference in smaller tetrahedral cavities [66], both selections being made in the presence of manipulated amounts of the free cations and limited production of proteins due to feedback. The free cation concentrations arise through differential pumping and buffering in different compartments while free protein is limited by regulation at the DNA.

An interesting example where Fe^{2+} is preferred over Zn^{2+} by a combination of binding strength and free ion concentration differences is in M^{2+} binding to Fe_3S_4 of redox enzymes [65]. To explain this observation we describe a new factor which can stabilise metal ions of a given kind in a cluster and then show how this effect can help to selectively incorporate Fe^{2+} in such a site in proteins. We shall consider the equilibrium (Fig. 29):

$$Fe_3S_4 + M \rightleftarrows Fe_3S_4M$$

where M can be any cation but especially we look at the ions, which could bind to S^{2-}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Zn^{2+} and Cd^{2+} as well as Cu^{+}, derived from Cu^{2+}. The equilibrium binding data are given in Table All the ions are in fast exchange except for Cu^{2+} which goes to Cu^{+}. We must comment on the rather

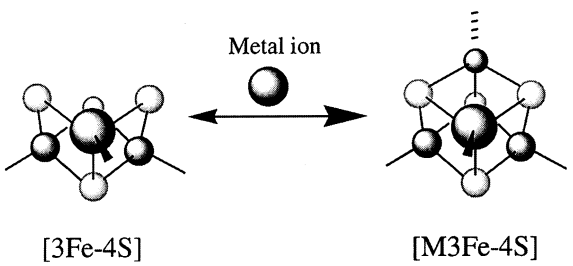


Fig. 29. The reversible exchange of iron (Fe^{2+}) in a Fe_4S_4 protein which can be used as a control [28,29] in Fnr or the aconitase transcription factors.

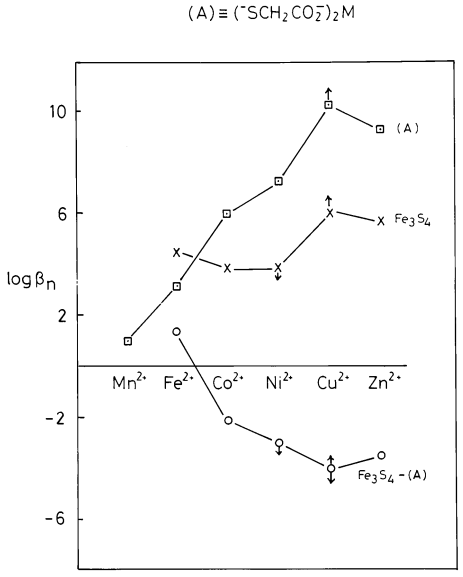


Fig. 30. The stability constants of various metals in the Fe_3S_4M cluster of a protein, and in (A) the chelate of bis-thioglycollate and the difference between them. Note how only Fe^{2+} binds to the cluster preferentially. Data from Ref. [23,65].

high value for Fe^{2+} relative to that of the other ions, Fig. 30, which is not expected from the Irving–Williams series, but first we must be sure that all the ions are in their usual electronic structures i.e. are in high-spin ground states.

The spin states of the Fe_3S_4 ($S = 2$) and the Fe_3S_4M centres are given in Table 26. The conclusions are simple. The spin state remains as in Fe_3S_4 for $M = Zn^{2+}$ and Cd^{2+} in Fe_3S_4M but becomes that of Cu^{+} on addition of Cu^{2+}. Antiferromagnetic coupling reduces the spin in the cases of Ni^{2+} ($S = 1$), Co^{2+} ($S = 3/2$) and Fe^{2+}

Table 26
Equilibrium constants K_d ($= 1/\beta$) for reaction (1) as determined for *Desulfovibrio africanus* ferredoxin III, and ground spin states (S) observed for $[M3Fe\text{-}4S]^{(n+2)+}$ cubane adducts formed in *Da*, *Pf* and *Dg* ferredoxins [23]

	Core	Adduct with-						
	$[3Fe\text{-}4S]^n$	Fe(II)	Co(II)	Ni(II)	Cu(I)[a]	Zn(II)	Cd(II)	Tl(I)
K_d (μM)		30	130	nd[b]	<1	1.6	0.8	1
S ($n = 0$)	2	0	1/2	?	2	2	2	2
S ($n = 1$)	1/2				1/2			1/2
S ($n = -1$)	5/2	1/2(3/2)	1			5/2	5/2	

[a] Reaction studied with Cu^{2+} in solution, but from spectroscopic data, Cu is bound as Cu(I).
[b] Formation of a Ni adduct in *Da* Fd III has not been detected, up to a Ni^{2+} concentration of 1 mM Ni^{2+} (Ref. [65,81]).

($S = 2$) but there is no change of local spin from ground state ions of any atom. There is an additional peculiarity of the Fe_3S_4Fe cluster however in that the four iron atoms are closely identical — a mixed valence state of class II — according to Robin and Day [68] arises. Now such a state has a stabilisation energy due to what is often called resonance in that electrons are spread over extended molecular orbitals of several identical atoms. We consider that this resonance energy helps to give an increased stability to the Fe_3S_4Fe cluster over that expected from the Irving–Williams series compare

$$Cu^{2+} > Zn^{2+} > Ni^{2+} > Co^{2+} > Fe^{2+} > Mn^{2+} \text{ (Irving–Williams)}$$

and

$$Cu^{2+} > Zn^{2+} > Fe^{2+} > Co^{2+} > Ni^{2+} > Mn^{2+} \text{ (for } Fe_3S_4)$$

However these data (Fig. 30) still show that the absolute binding order would lead preferentially to clusters of the kind Fe_3S_4Zn or Fe_3S_4Cu not Fe_3S_4Fe when metal ions are equally available. We shall now show why this does not happen in a cell due to competitive binding by other ligands than Fe_3S_4, which, as is explained in Section 4.3, controls the relative availability of free metal ions in the cytoplasm.

The first point to observe is that while *thermodynamic* binding constants to Fe_3S_4 are in the order given this order applies to conditions in which the free metal ions are in equal concentration. It is the *effective* binding ability under the buffering and pumping conditions of the cytoplasm which defines the true competition for this ligand in a cell. We cannot give exact numbers but we considered in Section 1.5 (see Fig. 2) that the concentrations in the cytoplasm of many cells could be as different as

Zn^{2+}	Ni^{2+}	Co^{2+}	Fe^{2+}	Mn^{2+}	Mg^{2+}
10^{-11} M	10^{-9} M	10^{-8} M	10^{-7} M	10^{-6} M	10^{-3} M

For ligands using three or more good donors and especially when two or more of the three or more donor atoms are N- or RS^--donors as in a cell cytoplasm. The thermodynamic binding constant difference between Zn^{2+} and Fe^{2+} exceeds 10^5 so that selection is in favour of Zn^{2+} and Fe^{2+} cannot bind to these ligands when they are produced only in limited quantitative concentrations opposite the available Zn^{2+}, Cu^{2+} (Cu^+) and Cd^{2+}. At the same time it follows that in the presence of these strong ligands, e.g. thiolates, the metal ions Cu^{2+}(Cu^+), Zn^{2+}, Cd^{2+} will be effectively removed from any competition for weaker binding centres such as Fe_3S_4.

The importance of this particular selection of relatively weak binding of Fe^{2+} in Fe/S clusters (and the preference over all other metal ions) is that it allows the *equilibration* of a low concentration of free Fe^{2+} from the Fe_4S_4 clusters of major transcription factors, Fnr in bacteria and aconitase in higher organisms, to control production of a series of redox enzymes (see Table 11). Hence Fe^{2+} is a major homeostatic redox device in cells. Notice that the binding constant to many Fe^{2+} enzymes in Table 11 is also around 10^7.

A parallel but equally effective control equilibrium is that between Fe^{2+} and the transcription factor Fur [31,32]. Here the weak exchangeable binding of Fe^{2+} $K \approx 10^7$ is again protected from exchangeable zinc due to the low availability, tight binding of Zn^{2+} in cells to N/S donors in tetrahedral array, see the protein Zur, Section 4.3. It is also protected from cobalt and nickel since these metal ions are transported by kinetically protected means, Sections 5.10 and 5.11, but not from free Mn^{2+}. It is quite probable that Fur, but not Fnr, also acts as an Mn^{2+} transcription factor but especially in organisms which use little iron, e.g. lactobacilli. Taken together Fur and Fnr, together with other transcription factors of very different kinds based on phosphorylation, regulate a vast range of redox enzyme production.

N.B. It is notable that the Fe_2S_2 clusters are all of low potential (-0.35 V) and iron is not known to exchange from them. In fact the iron is held in the resting state as Fe^{3+}. Fe_2S_2 clusters do not generate transcription factors.

6.3. The incorporation of Fe^{3+} into biological chelates

The difficulties of incorporating Fe^{3+} into soluble coordination compounds in organisms are two-fold: the high pH of cells and even animal extracellular fluids of around 7.5 and the low redox potential of the cytoplasm which may well approach -0.25 V. The pH reduces the free Fe^{3+} concentration to about 10^{-18} M however if the free Fe^{2+} concentration is only 10^{-8} M due to uptake into the cytoplasm of Fe^{2+} then using the cytoplasmic redox potential of -0.25 V the concentration of free Fe^{3+} must be much lower at about 10^{-24} M. These data apply to equilibrium conditions and it is in fact very probable that the free Fe^{3+} ion is in equilibrium with both the pH and the free Fe^{2+} in the cytoplasm due to the presence of exchangeable and redox reversible catalytic centres for electron transfer. This does not mean that all bound states of Fe^{3+} are also in equilibrium. However as stated before it is easier to analyse incorporation at equilibrium than in sites occupied by kinetically managed insertion.

In essence to hold iron in solution at pH 7.5 in an oxidising medium, e.g. the aerobic environment or the extracellular fluids of animals (not plants), a ligand with a binding constant greater than 10^{18} M had to be devised. The values for certain

Table 27
Some iron stability constants and redox potentials at pH 7 [23]

Ligand	log K(III)	log K(II)[a]	$E°$ (pH 7.5)
Transferrin	21.5	4.0	−0.36
Ferri-oxamin B	21.8	7.2	−0.17
EDTA	25.1	14.3	+0.05
8-OH-quinoline ($\log\beta_3$)	38.0	22.1	−0.26
Rubredoxin	26.0	8.0	−0.35

[a] N.B. (1) The free Fe^{2+} ion concentration in cells is 10^{-7} to 10^{-8} M. (2) In lysozomes the pH 5 when the redox potential rises and Fe^{2+} is released.

biological chelating agents exceed this [23] Table 27. Note that these ligands do not retain Fe^{2+} very strongly whence on entering the cell, e.g. in vesicles at a low redox potential (and low pH) Fe^{3+} is reduced and Fe^{2+} released for further processing.

To reconvert the Fe^{2+} ion to Fe^{3+} after it reaches the cytoplasm requires the redox potential of around -0.25 V to be taken into account. As stated above given that free Fe^{2+} is about 10^{-8} M this requires that free Fe^{3+} will be no higher than 10^{-24} M in the solution so that a chelating agent with a stability constant for its Fe^{3+} complex of $>10^{24}$ M is now required. No good values are available but the very nature of the ways in which Fe^{3+} ions are bound in proteins suggests that this value can be easily exceeded. In rubredoxin the value of log K is estimated to be 26, see below [70].

It is always possible in an aerobic cell that given a suitable enzyme which is sensitive to dioxygen, local conditions in a vesicle or small container will be very different from those in the cytoplasm and no longer dominated by glutathione redox potentials of -0.25 V. The appearance of ferritin, effectively $Fe(OH)_3$ in the cytoplasm of many cells, which would equilibrate with a free Fe^{3+} concentration of 10^{-18} M is therefore intriguing. The obvious conclusion is that this enclosed particle does not equilibrate with Fe^{3+} ions in the cytoplasm. Release of iron must then be due to reduction. In fact most if not all of the reactions of free iron in cells are reactions of Fe^{2+}.

Note that if a cell had used ascorbic acid at pH 7 to maintain cytoplasmic redox potentials at $+0.1$ V then the standing Fe^{3+} concentration opposite $Fe(OH)_3$ of 10^{-18} M and a free Fe^{2+} concentration of 10^{-8} M would match this redox potential almost exactly. Ascorbic acid is a major reducing agent *outside* animal cells and both outside and inside plant cells. In plants ascorbate in the extracellular fluid is at a pH 5.5 which allows even higher levels of free Fe^{3+}. In fact this pH allows a $[Fe^{3+}]$ of 10^{-12}M. The binding constant of citrate at this pH is around 10^{12} so that citrate can carry Fe^{3+} iron in external fluids of plants but not of animals. The binding of citrate to Fe^{2+} has a log K of 10^3 so that the redox potential is about $+0.15$ V close to that of ascorbate and the bound Fe^{3+} is then not greatly reduced by this vitamin.

An interesting peculiarity is the isolation of rubredoxin with four RS^- donor centres binding Fe^{3+} while the same four centres bind Zn^{2+} in zinc fingers [69]. Again the same four RS^- groups in metallo-thionein bind Zn^{2+} or Cu^+ but this protein in cells does not bind Fe^{3+} (or Fe^{2+}). It is easily seen why rubredoxin is extracted with Fe^{3+}, not Fe^{2+} and not Zn^{2+}. Experiments show that the Fe^{3+} state is in fact some 10^{17} times more stable than that of Fe^{2+}, and the redox potential of rubredoxin is -0.35 V which is below the potential of the redox buffer, glutathione. The protein is then a Fe^{3+} not a Fe^{2+} protein in its resting state. The binding constant of Fe^{3+} to rubredoxin is known to be 10^9 times stronger than that for Zn^{2+} so that Zn^{2+} binds 10^8 times more strongly than Fe^{2+} — exactly as expected from model data on B.A.L. Fig. 17. Given these differences and the binding of Zn^{2+} to other types of centre it is clear why rubredoxin binds Fe^{3+} not Zn^{2+}.

The zinc finger and metallo-thionein proteins are constructed quite differently from rubredoxin in that the former is a strongly folded β-sheet (even as an apoprotein) while the other two proteins have little conventional secondary structure and are unfolded as apoproteins. Thus rubredoxin forms a much tighter fold which gives shorter M–S bond distances and hence stabilises Fe^{3+} relative to Fe^{2+} or Zn^{2+}. The suggestion is that Fe^{3+} binding is much reduced in zinc fingers or metallo-thioneins by the longer M–S bonds so that Fe^{2+} (and more so Zn^{2+}) is favoured and the redox potential elevated. An estimate of the Zn^{2+} binding constant to rubredoxin is 10^{17} while that of zinc metallo-thionein and zinc fingers is around 10^{11} so that Zn^{2+} binding to rubredoxin is effectively blocked by Fe^{3+}. Thus it is cavity size which decides selectivity under conditions of competitive binding by the same group of donors.

This coordination group of four RS^- donors provides an excellent example of the fact that different protein folds can so control metal bonding that selectivity using the same donors can give rise to different stability orders and different redox potentials. It is also observed that through a change of the primary sequence of the proteins but using the same four donors, here four RS^- groups, the geometry around the metal ion can be adjusted. [69] While expectation for both Fe^{3+} and Zn^{2+} is that the geometry should be tetrahedral, this is very closely the case in some rubredoxins and zinc fingers but far from so for the centre of desulforedoxin which has the same donors. Thus bond angles as well as bond distances can be used to affect stability — both depending on the protein fold — in constrained metal ion states.

6.4. The ferric sulfides and other metal ion clusters [24]

Now we also have to explain the fact that iron clusters such as Fe_2S_2, Fe_3S_4 and Fe_4S_4 exist in proteins while sulfide clusters of the other elements are absent. There are however clusters of M_2O_2 or $M_2(OH)_2$ in many enzymes and we have already met clusters of $M_2(RCO_2^-)$ or mixed clusters of formulae $M_2(RCO_2^-, OH^-)$ in non-redox enzymes. In the case of simple O-donors, O^{2-} or OH^- clusters are easily produced in the M^{3+} state for both Fe and Mn in oxidising conditions even in the absence of proteins, but other metals do not easily occur as M^{3+} ions. The stability of these particular M^{3+} clusters is due to the ease of oxidation of M^{2+} ions in sites of very small anions of high charge density, which bind M^{3+} better than M^{2+}. When the ligand size is increased and charge density is reduced we observe the production of clusters $M_2(RCO_2^-\ OH^-)$ and $M_2(RCO_2^-)_2$ for divalent ions, Zn^{2+}, Mg^{2+}, Ni^{2+}, Fe^{2+} Mn^{2+} and Co^{2+} rather than for trivalent cations, due to the reduced binding of M^{3+} relative to M^{2+} to RCO_2^- compared with that to OH^- or O^{2-} which are stronger bases. However introduction of tyrosine RO^- ligands again allows M^{3+} proteins to form. Note that divalent metal ions do not generate basic carbonates generally at pH 7 and a contributory cause of the M_2 (RCO_2^-, OH^-) cluster is the peculiarity of the protein fold.

Now the peculiarity of iron rests in the ease with which it is oxidised to Fe^{3+} when bound by anions even in the presence of reducing ligands such as RS^- and

S^{2-}. The presence of M/S clusters using Fe and no other metal is due therefore to the stabilising influence of Fe^{3+}. Thus nature has taken advantage of the low ionisation potential to the M^{3+} state of iron relative to that for any other common metal ion.

This is best seen by reference to mineral chemistry. A well known mineral which is formed in reducing conditions (presence of sulfide or H_2S) is chalcopyrite $CuFeS_2$. This mineral is not a member of the pyrite family (Fe^{2+} S_2^{2-}) but should be written Fe^{3+}, Cu^{+} $(S^{2-})_2$ when in local regions it resembles the cluster of ferredoxins Cu^{+} Fe_3^{3+} $(S^{2-})_4$. Thus in the presence of sulfide the formation of Fe^{3+} containing cluster complexes occurs readily even when copper is reduced to Cu^{+}. The studies of many model inorganic clusters shows that Fe^{2+}/Fe^{3+} mixed sulfide clusters are stable but all Fe^{2+} cluster are not. In some bacteria a magnetic mineral Fe_3S_4, covellite, is formed in vesicles.

Similar problems arise in the stabilisation of higher oxidation states of manganese and cobalt. In fact Mn^{3+} is the stable form of a superoxide dismutase in air but given the high redox potential of the manganese protein of around $+0.3$ V, the Mn^{2+} state is present inside cells. In the three ribonucleotide reductases a low redox potential is reached in the strong binding complexes of Fe_2O^{4+}, Mn_2O^{4+} and cobalamins, all of which are in trivalent metal ions states. Bringing such higher oxidation states as $+3$ into the cytoplasmic range of redox potentials, below 0.0 V, is of the greatest difficulty without using special complexing situations such as these. However such states are known for nitrile hydratase which holds either Co^{3+} or Fe^{3+} in low-spin states using binding to five anionic RS^{-} donors and an ionised peptide N^{-}group. The still higher oxidation states of manganese Mn^{4+}, and iron Fe^{4+} are only available in non-equilibrated excited states, e.g. in O_2 production and use. N.B. M^{4+} is better written MO^{2+}.

Of course higher oxidation states are easily accessible for metals early in the transition metal series which even form oxyanions, e.g. VO_4^{3-} and MoO_4^{2-}.

6.5. Copper incorporation [70]

Once copper as copper(II) became available in the environment (see Section 8) the problems of its incorporation were two-fold: (1) its binding to ligands is extremely strong and in particular Cu^{2+} and especially Cu^{+} bind to N- or S-donors much more strongly than Zn^{2+} which itself binds the most strongly of the other divalent ions, Section 3; and (2) protection had to be found against copper as an oxidising catalyst in the cytoplasm The solution to these problems of selective binding in the reducing cytoplasm lay in using the stereochemical properties, covalence and size of Cu^{+}. This state Cu^{+} is general in the cell cytoplasm due to the fact that the redox potential of copper in free and bound forms is always greater than $+0.1$ V [1]. The peculiarity of Cu^{+}, shared only by unavailable ions such as Hg^{2+} and Ag^{+} is that it has a very high affinity for thiolate ligands in a *linear* 2-coordinate or a three coordinate trigonal structure. Cu^{+} can be removed by such sites from competition while allowing zinc and iron to bind to RS^{-} and S^{2-} (in tetrahedra and their clusters) as described above. In fact a very similar sequence in

metallo-thionein binds both Cu and Zn cooperatively but in different structures. This selective activity only operates so long as excess apoprotein for the binding of copper is not synthesised and excess copper over the binding proteins does not occur due to the selective pumping of the Cu^+ into vesicles or out of cells by proteins containing two linear RS^- donor centres. Due to the specificity of this chemistry associated with protein thiolate groups free copper (Cu^+) in the cytoplasm is tightly controlled by its own reversible transcription factors to below a nominal value of 10^{-15}M.

Once expelled outside the cell the copper had to be handled differently since it can be stable there as Cu^{2+}. Here the strength of binding to N- or S-donors is more powerful than that for any other commonly available biological metal ion. However these centres must only be produced in limited amounts and other metal ions had to be bound to alternative sites of greater stability for them since otherwise excess of copper apoproteins could pick up several metal ions, e.g. Zn^{2+}. Three possible incorporation modes for securing highly selective binding of copper are used: (1) binding to exchanging sites in the cytoplasm of pumps, buffers and transcription factors with uniquely strong binding for copper as Cu^+; (2) moving the copper as Cu^+ into protected environments, vesicle compartments, before it is incorporated and oxidised even in reversible sites (it does appear that copper proteins and enzymes pass through vesicle routes different from those for say zinc). (3) At the final state of incorporation in a protein the copper is *irreversibly* locked in a site so that exchange and competition are not problems.

One interesting example of (2) occurs in eukaryotic Zn/Cu superoxide dismutase. Here the incorporation of both metals in the cytoplasm is in the final state irreversible but their individual sites have to be filled sequentially and selectively. Both sites are formed from histidine N-donors. It is probable that the zinc site is too weak to bind copper given the buffering by metallo-thioneins. This site is

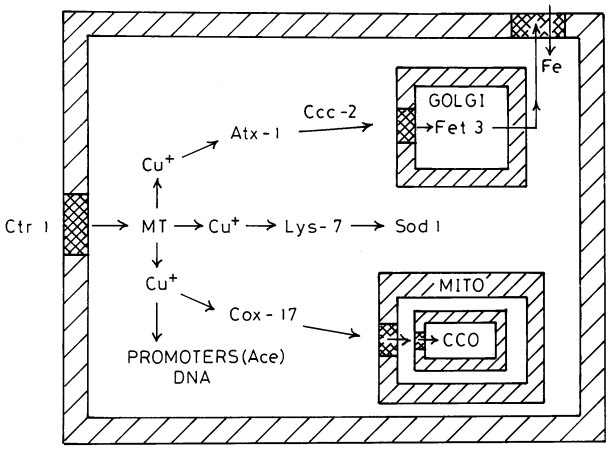


Fig. 31. The irreversible distribution of copper except for the slow exchange with transcription factors such as ACE and with the storage protein, metallo-thionein [70]. The free Ca^+ is nominal.

tetrahedral. Once the zinc binds it folds the protein more tightly and induces the second site for copper which now has an ionised imidazole, N^--donor. The site is not far from tetragonal. The binding site has a constant for copper approaching 10^{15}. Thus it is the binding of zinc possibly with a constant of 10^{11} which makes the binding of copper possible and then both are irreversibly held. There is evidence that the copper is transported to the site by a specific copper transport protein [70]. The specific ways of handling copper by such carriers are shown in Fig. 31 but we do not know the time constants for copper exchange.

6.6. Molybdenum enzymes [15,27,71]

A *minor* element in redox actions of proteins is molybdenum. We have described its uptake in Section 3. The metal does not equilibrate from its co-enzyme form as a dithiolate. The selective uptake may be a purely kinetic handling of molybdate, or in parallel cases of tungstate, so that competition for the site is reduced. However the total production of the dithiolate must not exceed the availability of molybdenum since this is a powerful ligand which could bind to Fe^{3+}, or some of the divalent ions. In the absence of equilibrium data for model systems we do not know how severe competition could be.

The molybdenum acts as a two electron (O- or H^--donor/acceptor) and it does not equilibrate electrons with other redox couples except those directly involved in enzymic activity.

Further elements of importance in redox reactions are nickel, [72] cobalt, in vitamin B_{12} (see Sections 7 and 8) and manganese. The use of nickel [73] is slight but we refer to it again in the complexed form of F-430 in Sections 7 and 8 while the use of manganese in redox reactions except in the case of O_2-generation is very rare. Little is known as yet about the manganese coordination in redox enzymes.

6.7. Conclusion concerning redox metal ions

In summary redox active metals are frequently in non-exchanging environments so that their distribution both as to preferred binding sites in proteins and the selective positioning of these proteins is difficult to analyse. The case of molybdenum, Section 3.5, illustrates the problem. The exceptions would appear to be ferrous iron, Fe^{2+}, and possibly Mn^{2+}, which exert control and regulatory functions since they exchange easily. These functions appear to be extremely primitive.

7. Kinetic insertion mechanisms [74]

As already described, a way in which thermodynamic competition is avoided in the cytoplasm and especially at the final site of placement of a metal ion in a site is for the ion to be selected at an earlier stage and then for it to be inserted irreversibly into the required site (see Sections 4–6). This is the mechanism for

insertion of Mg^{2+} into chlorins to make chlorophyll, Fig. 32. N.B. It is also the general method for incorporation of non-metals (see Section 2). For Mg^{2+} the initial step could be simple thermodynamic binding to a weak binding centre, such as ATP, since all other metal ions do not bind this ligand because their binding constants to ATP fall below 10^{-5} M and their free concentrations are much less than 10^{-5} M (see Figs. 17 and 18). Only Mg^{2+} is present at high enough concentrations, 10^{-3} M, to bind. At the other extreme some centres can select thermodynamically Cu^{+} or Fe^{3+} preferentially to such a degree that in the competition between metal ions and ligands for binding sites these metal ions have virtually specific carriers provided for example by two linear RS^{-} groups for Cu^{+} and in part by tyrosine anions for Fe^{3+}. The strength of binding is such that the combination is effectively irreversible. Now this leaves the metal ions which bind with intermediate strength Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+} and Zn^{2+} in competition for mixed N-donor and O-donor sites. A simple solution for generating selectivity, as described above, is that environment factors which determine relative availability together with the thermodynamics of binding decide selection and hence a given protein may have a site of sufficiently good selectivity without further assistance. In fact quite a number of metal/protein combinations fall into this group, Sections 4–6. However there is a further group of proteins which seem to bind only a particular, probably a unique, cation, despite a coordination sphere which does not appear to offer advantageous thermodynamic binding. A striking example is urease

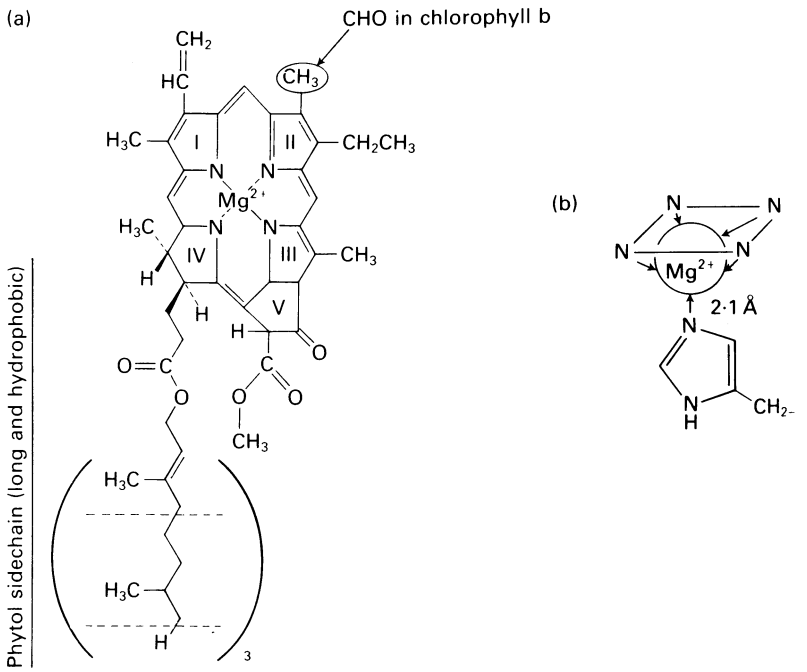


Fig. 32. The irreversible coordination of Mg^{2+} in chlorophyll.

Table 28
Irreversible pumps and proteins for cations

Cation	Pump[a]	Transcription factor or Carrier[b]	(Protein) product enzyme
Co^{2+}	ATP (O), ER	Cob(D)	Calreticulum (ER)
Ni^{2+}	ATPase (H)	Hox N, Hup N Hyp, Hyc	Hydrogenase
		Ure H, Nix A	Urease
		CooH	Carbon monoxide Hydrogenase
Co^{2+}	ATPase (H,S)	nhlF	Nitrile hydratase
		Cot-1	Mitochondrial transport?
		Chelatase(Co)	Vitamin B_{12}
Cu^{2+}/Cu^{+}	ATPase (S)	Lys-7	Superoxide dismutase
		Cox-17	Cytochrome *c* oxidase
		Atx-1	Vesicle enzymes (hydroxylases)
Zn^{2+}	ATPase (H,S)	?	Vesicle enzymes (proteases)
Fe^{3+}/Fe^{2+}	see Fe^{3+} (H,O) (transferrin receptor)	Chelatase (Fe)?	Various heme proteins (Cytochromes)
Mg^{2+}	ATPase (O)	Chelatase (Mg)?	Various Chlorophyll
Ni^{2+}	ATPase (H)	Chelatase (Ni)	Proteins (Reaction Centre) F-430

[a] The different ATP-ases have very different selective binding donor atoms, O (carboxylate), H (histidine), or S (cysteine), and stereochemistries. They may equilibrate with two different metal-ion concentrations on two different sides of a membrane.

[b] The nomenclature in this column often also refers to a gene product, e.g. Cox is a protein product of the Cytochrome Oxidase gene, *Cox*.

which contains nickel in two sites both of which appear unsuitable for it and into which zinc would fit quite well (see Table 16). The evidence is now strongly in favour of the idea that this situation has arisen through *a selective insertion* of Ni^{2+} into the site. The steps in the process may have been generated quite early in evolution since this enzyme is found in primitive bacteria and they could involve a series of handing-on reactions. Starting at the membrane surface of the cell and the Ni^{2+} ion on the outside the first step is a binding to a Ni^{2+} pump. Binding is competitive and reversible but the Ni^{2+} binding could well generate a particular geometry internal to the pump, say an octahedron of given dimensions and donors, and only if this geometry prevails is an ion, here Ni^{2+}, pumped toward the cytoplasm. The pumping costs energy from ATP. The Ni^{2+} is thereafter handed on directly to a specific transport protein in the cytoplasm which accepts the ion from the pump with the result that the cation is never free. Proteins of this kind are known and seem to be able to deliver the Ni^{2+} to urease, Section 5.10. We then believe that this is an example of a more general case where pumps and carriers deliver a selected metal ion to a specific protein. Table 28 lists some different systems for elements where such kinetic selectivity often appears to start from the internal (cytoplasmic) side of pumps. Note that iron and manganese may not use such insertion mechanisms into proteins generally since they have control functions

in cells and must therefore exchange from many sites, excluding heme of course. NB. Units such as heme act as controls in cells quite separately from the metal ion in them.

Quite interestingly as stated earlier, Section 4.1, many of the zinc, nickel and cobalt carrier proteins have several histidine residues [1]. In the extracellular fluids of multicellular animals and plants transport earlier than that in the cytoplasm is needed and we find generally that histidine residues act at the carrier site [38]. Ni(histidine) is the known carrier in hyper-accumulator plants (which also accumulate Co^{2+}) [39], while in animals albumins carry zinc and copper using histidine side chains. It is these carriers which must be recognised by the selective pumps shown in Table 28. In fact the Cu^{2+} albumin carrier is known to utilise peptide $>N^-$ anion centres for selective binding which will not bind other cations at pH 7.

We turn finally to the way in which all the proteins for selecting metal ions have evolved so that they may perform different kinetic, not thermodynamic, functions in the exchange and retention of metal ions.

7.1. Protein folds and kinetics

If a metal ion such as zinc or copper is to be in relatively fast exchange then it cannot be bound by four (N,S) donors in the interior of a protein with a strong tertiary fold since protein unfolding with loss of strong binding would be too slow. Even if the protein has three-donor groups and the metal ion is held in a strong fold, exchange can be very restricted. Hence it is useful to look at the folds of the proteins holding metal ions to discover how such metal ions can be effectively in fast exchange. We find that many metal-binding proteins which do not allow exchange have a multi-stranded β-sheet construction. In these cases the metal ion is locked into a constrained site such as that of azurin (for copper), and the binding can be considered to be irreversible. Notice how very different the non-exchanging trapping of zinc and copper is from that of magnesium in such compounds as chlorophyll. At the other extreme are proteins which only fold when bound by metal ions. Typical proteins are metallo-thionein, binding zinc and copper, and zinc-fingers. The fold of these small proteins is not readily described in simple terms. Their apoproteins, which have 'random coil' structures and weak folds may then equilibrate when folding with free metal ions according to simple unassisted on/off rate constants when the k_{off}/k_{on} gives the binding constant, K. If k_{on} approaches the diffusion limit of say 10^9, as it can for zinc and copper binding to readily available protein surface sites, then the off-rate governs exchange. Consider a binding constant of 10^{12} M^{-1} which is close to the value for zinc, then given $k_{on} = 10^9$ the off-time constant is 10^3 s which is about 20 min. For binding constants of 10^{11} M^{-1} it is only 20 min so that zinc concentration changes can readily act as regulatory switches for a slow process. Although they are therefore of little use to bacteria, doubling time 30 min, they are increasingly valuable as the complexity and life-time of an organism increases i.e. in eukaryotes and then multicellular organisms. In fact Table 30 shows that the number of zinc (and copper) proteins increases vastly in these organisms as life evolved from bacteria and even to this day bacteria are low in zinc (and copper).

There are also those proteins which fold in α-helices. While they do not unfold readily they do have adjustable structures so that they give rise to pumps, exchangers, switches and gates. This ability to adjust, open and close, can allow much faster exchange of bound metal ions (as is necessary in these devices) than does a β-sheet construct. Examples are given in References [1] and [2]. Particularly note the helical calcium proteins which are internal cellular triggers where exchange rates are 10^{-3} s even though the binding constant is 10^{6} M^{-1}.

Overall we see that the dynamics of the protein fold can control the exchange of metal ions and hence an organism may utilise the fold kinetics in the selection of metal ions and their functions.

8. Evolution of selection

8.1. Historical development of metal distribution: introduction

Examination of different organisms and especially looking at the differences between anaerobic, presumably more primitive organisms, and aerobic life-forms indicates that there have been very substantial changes in the way elements are distributed in cells [2,3]. To appreciate these developments we note first the changes in the chemistry of the environment of the Earth's surface [2]. Starting from a strongly reducing atmosphere, probably containing gases such as H_2, CH_4, NH_3, H_2S and H_2Se as well as some more oxidised products such as CO/CO_2 and N_2 all in the presence of water (H_2O), the atmosphere switched over 3.5 billion years to the present day N_2, O_2, CO_2 air around us. At the same time surface land has developed from being in part sulfides to being almost totally oxides. The early reducing conditions of the sea with a redox potential of around 0.0 V or below were probably approximately constant for some one or so billion years and given that the temperature was quite high we can assume that the mineral content of the sea was fixed in thermodynamic equilibrium with the early atmosphere and the early earth. The availability of certain elements was then constrained mostly by the presence of sulfide. We can be confident that there was virtually no Hg^{2+}, Pb^{2+}, Cu^{2+}/Cu^{+}, and Cd^{2+} in the sea and that there was rather little Zn^{2+} and only just a little more Ni^{2+} and Co^{2+}. Fe^{2+} and Mn^{2+} were more plentiful, while Mg^{2+} and Ca^{2+} were relatively freely available though the Ca^{2+} was buffered by carbonate formation. Iron was able to form some soluble clusters with sulfide, $Fe^{2+}/Fe^{3+}/S^{2-}$, in which some Fe could be replaced by Ni and some of the S^{2-} could be replaced by Se^{2-}. Larger anionic units such as MoS_4^{2-} and VS_3^{-} were also able to bind to the clusters (see [24]).

Of the metal elements which do not form sulfide precipitates or complexes those which gave more soluble hydroxides and carbonates and were of considerable abundance were readily available as cations especially Na^{+}, K^{+}, Mg^{2+} and Ca^{2+}. All other cations were of low availability either due to low abundance, Li^{+}, Be^{2+}, boron, and heavy elements of many Groups of the Periodic Table or to insolubility of their oxides or hydroxides, e.g. of Al^{3+}, Cr^{3+} and of ions of Group 3, 4, 13, and

14. The lighter non-metals such as H, C, N, O, S, Cl, and Se were all available as simple substances such as H_2O (H_2S), CO, NH_3, OH_2, SH_2 and SeH_2, or as simple or more complex anions such as Cl^- and HPO_4^{2-}. Notice that boron and many heavy non-metals were largely absent due to their low abundances while there were present considerable quantities of $F^-$and $Si(OH)_4$. All elements which were available and could be directly covalently bound under primitive conditions, H, C, N, O, S, and Se, were incorporated into *coded* amino acids.

In these circumstances the further binding in cells especially of Fe, Ni, V and Mo whether in clusters or not was to RS^- and RSe^- as well as N-donor side-chains of proteins. Some zinc was also bound to these RS^- and N-donors but not in clusters. The common metal elements in the cytoplasm of anaerobes and bound to cysteine even today are Fe^{2+}, Zn^{2+}, Ni^{2+} and Mo. In any initial circumstances of biological consequence elements such as Na^+, K^+, Mg^{2+}, Ca^{2+} and Mn^{2+} could not have been bound to sulfur groups, and would be bound overwhelmingly to oxygen, O, donors. In these early conditions zinc could be captured only with great difficulty if at all by N- and O-sites due to preferential removal of the metal by N-, S-donors or by binding to sulfide. This thermodynamic analysis indicates that N-donors or combinations of N- and O-donors in the cytoplasm would be left to bind some Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+} and possibly Mg^{2+}, as we have seen in Sections 4 and 5. None of this discussion of thermodynamics excludes the possibility that quite other binding sites for various metals were generated here and there by special stereochemical combinations or by kinetic traps, e.g. porphyrins, but many of them could only have developed later (see Section 8.4), i.e. after the most primitive selection by simple donors.

Given the known composition of today's anaerobic cells, thought to be related to primitive cells which contain not just sodium, potassium, magnesium, iron and manganese, but some cobalt, nickel, molybdenum and zinc, we are forced to conclude, in agreement with the general gist of the above discussion, that all these metal elements but not copper or cadmium were able to be absorbed to some extent by early life, Table 29. Zinc was used at first, as far as we can tell, not so much as a catalyst, where it has to be coordinatively unsaturated, but as a strongly bound structure-forming unit in the primitive DNA and RNA polymerases.

Table 29
Original anaerobic competition (sulfide media)[a]

	$Mn^{2+}(10^{-6}$ M)	$Fe^{2+}(10^{-7}$ M)	$Zn^{2+}(10^{-11}$ M)
Ligands	Mainly-CO_2^- → and 1-2 N →	Mainly RS^-(clusters) and some N/O →	Mainly N → or RS^- donors
Geometry	6-coordinate	4-coordinate clusters (S^{2-}) six-coordinate (N/O →)	4-coordinate
Exchange	Fast $>10^{-1}$ s	Fast $>10^{-1}$ s	Slow $<10^{+2}$ s

[a] The exchange rates meant that free iron (or manganese) could act as signalling ions to DNA-bound proteins of the $Fe^{2+}/Fe^{3+}/S^{2-}$ type but that Zn^{2+} could not be useful in signalling by exchange within the life-time of bacteria in fast growth (doubling time 30 min).

Note that these enzymes are highly structured even in the absence of zinc (see Table 19) and provide N/S-donor binding. The binding constant of zinc in them had to be somewhat in excess of 10^{11} M^{-1} to overcome the problem of the insolubility of zinc sulfide. Other metal ions such as Mn^{2+}, Fe^{2+}, Co^{2+} and Ni^{2+}, which, as stated, were more available although less powerful as Lewis acids, may well have been selected in such acid/base enzymes as nickel in urease, cobalt in methionine peptidase, iron (Fe^{2+}) in several systems, and manganese in prolidase; all use N/O-donor binding. There is no chemical reasoning which would suggest that freely available zinc could not (1) bind more strongly in the enzyme sites of these proteins and then (2) catalyse more effectively their hydrolytic reactions. It was probably just the lack of availability of zinc together with its affinity for N/S donor sites which lowered its concentration as a free ion and permitted selection of Fe^{2+}, Co^{2+} and Ni^{2+}. An indication of the fact that even cobalt and nickel were in quite short supply is that both were and are still frequently found in kinetically trapped environments, vitamin B_{12} and F-430 co-enzymes, rather than being directly bound to proteins and so exposed to loss of the metal ion by exchange. The sequestration of iron by heme, which came later also prevented loss. The binding or iron in Fe/S clusters does leave the metal exposed for further reactions so that heme iron is a much safer catalytic site. The use of these organic ring structures allowed the metal ions to be reactive while being retained. Apart from their functions in catalysis fairly rapid exchange, together with modest binding, allowed the elements Mg^{2+}, Fe^{2+}, Mn^{2+} (and much less if at all Co^{2+} and Ni^{2+}, which exchange more slowly) to be used as internal cytoplasmic messengers but at this stage Zn^{2+} could not function in this way since tight binding prevented exchange within the life cycle of bacteria (Table 29). Do not forget that to be useful in control exchange in a cell of half-life 30 min must be of the order of one minute or less.

8.2. The coming of dioxygen [2,3]

The very slow change from anaerobic to aerobic conditions on the surface of the earth allowed a gradual change in the availability of the elements [2,3]. If we assume that the progress of change was indeed very slow, over two billion years, from 1.5 to around 3.0 billion years ago it is reasonable to assume also that equilibrium between the sea and the surface soils and rocks was achieved all the time. We may then use redox potentials as a guide to change remembering that for some elements a change in redox potential or prevalent state may mean a reduction in availability while for others it will mean an increase. Elements which slowly became more available were Zn^{2+} (Ni^{2+} and Co^{2+}) followed later by copper as Cu^{2+}. Elements which became less available quite early were iron, precipitated as $Fe(OH)_3$, and sulfur as sulfide and selenium as selenide, which became sulfate and selenate. Some elements hardly changed in availability yet the possibility of change of oxidation state in their reactions led to them being activated by redox enzymes, e.g. peroxidases. Examples are the halides in order of ease of oxidation: $I^- > Br^- > Cl^-$ which could become XO^- or X^+, and so covalently incorporated in a variety of organic compounds.

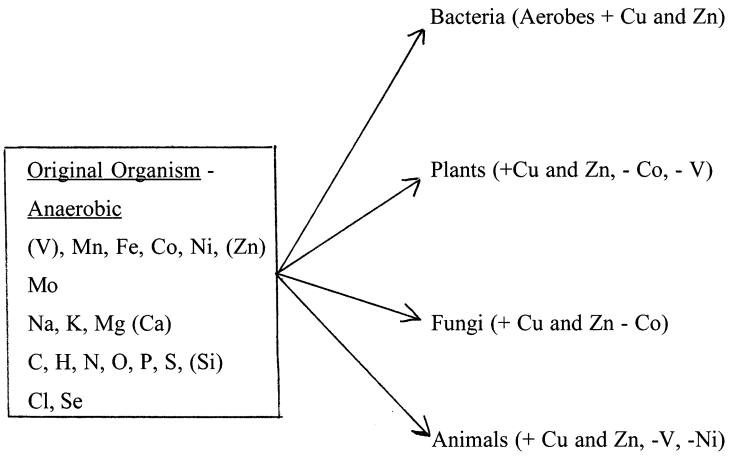


Fig. 33. The evolutionary relationships of the distribution of the elements. Note the roles of Co, Ni, Cu and Zn all of which have rather insoluble sulfides.

Increased availability of an element, A, might mean that, although a previously used element, B, was also increased in availability, the element A was now the better for a particular function. Thus we observe that in acid/base reactions zinc became the major catalyst while nickel, cobalt and ferrous iron all of which are generally less useful as Lewis acids, became less employed, Fig. 33. In fact cobalt is not used by higher plants and no nickel proteins are coded in the genome of higher animals. Methylation reactions dependent on vitamin B_{12} in prokaryotes are dependent on zinc enzymes in higher plants. Where the metabolism of early anaerobic prokaryotes using either cobalt or nickel has been kept in aerobes, protection of their binding sites from zinc is necessary. In fact we have seen that both cobalt and nickel have protected kinetic routes of uptake in those aerobic organisms that still use them, Sections 5.10 and 5.11. The greatest change due to redox potential changes of the environment is the great increase in the use of both zinc and copper.

Now evolution also led to increase in both size and complexity of organisms and they therefore have a considerably longer life-time. It was then necessary to develop new message systems [3] (see Table 30). First came the use of calcium, which does not depend on oxidative changes, but with the further development of organs in multicellular organisms, further new messengers were necessary outside cells. We see this in the development of a new class of cytoplasmic zinc proteins, the zinc fingers, which differ from the primitive zinc prokaryote cytoplasmic proteins in that, although they bind through N/S donors, the proteins allow faster exchange of zinc, now well within the life-time of the cell. As stated in Section 7 the difference between the proteins lies in the low stability of the apoproteins of the zinc fingers compared to that of the zinc polymerases of the prokaryotes. Zinc fingers apoproteins are virtually random coils. These zinc proteins signal the levels of zinc and of a variety of hormones, organic molecules, produced by oxidation, e.g. sterols, and distributed in the extracellular fluids from glands to organs.

Evolution also gave rise to the production of vesicles and some of these new compartments became used in the export of especially zinc hydrolytic enzymes (and copper oxidases later). The zinc in these proteins does not exchange but it is not held by N/S donors possibly due to the risk of extracellular oxidation of RS^- centres of the apoproteins since vesicles and extracellular fluids are exposed to high redox potentials. It would appear that the zinc is incorporated in these proteins inside the vesicles perhaps to avoid competition from Fe^{2+}, Co^{2+} and Ni^{2+}. The proteins have strong β-sheet folds even as apoezymes.

A picture of the evolution of calcium and zinc proteins is emerging as more and more genome sequences are becoming available, Table 30 [75,76].

It is not just the signalling by calcium and zinc which changed. Even prokaryotes which had used Fe/S proteins, through Fe^{2+} exchange, as transcription factors started using and use today additionally the Fur protein (N/O-binding) to act as a signalling transcription factor in aerobes. It may well be that the zinc uptake transcription factor, Zur, developed at roughly the same time. Once again note that these cytoplasmic proteins use N/S strong binding for zinc but N/O weaker binding for iron to avoid competition between the metal ions. Free iron signals to the genome at levels around 10^{-7} M while free zinc only acts as a signal at much lower level at around 10^{-11} M. Heme iron also became involved in new signalling due in part to the production of nitric oxide, NO, as a messenger. (N.B. NO binds to heme iron).

In the aerobic atmosphere the loss of substrates such as CO and H_2 have made much of the primitive functions of nickel and cobalt redundant in redox chemistry of aerobes while iron has found considerable new use in high redox potential chemistry of O_2, H_2O_2, NO etc. since FeO(N-ligated) states could be synthesised and are of functional value. However outside the cell, new oxidising substrates such as O_2 and H_2O_2 are handled more easily by copper catalysts than by iron or even heme iron since Fe^{2+} and heme itself are open to oxidative destruction. As stated above copper became available only after sulfide was largely removed from the environment [2,3].

Table 30
Metallo-protein sets in two organisms[a]

Protein set	Yeast	Worm (C. elegans)	Metal
Nuclear hormone receptor (Zn)[b]	0	270	Zn
Binuclear GAL Cluster (Zn)[b]	54	0	Zn
Metallo-proteases	0	94	Zn
Na^+ channels[b]	0	28	Na
Mg^{2+} adhesion[b]	4	43	Mg
Calmodulin-like proteins[b]	4	36	Ca
K^+-channels (voltage gated)[b]	1	68	K
EGF, Ca^{2+}-binding Cys-rich repeats[b]	0	135	Ca

[a] See Ref. [75].
[b] Absent in bacteria.

As described in Section 6.5 the functional importance of copper derives from the extremely strong binding of both Cu^{2+} and Cu^{+}, stronger for Cu^{+}, to both N and S donor sites of proteins. This allows the metal/proteins to be exported but it also gives them a relatively high redox potential even matching that of O_2 of the atmosphere. It was then a natural development to use copper, once it became available, as the oxidation centre outside cells. In particular copper became used in oxidative catalysis to bring about cross-linking stabilisation of connective tissue. In this way cells and organs could become organised into multicellular structure. There is also a quite different demand for growth-the breakdown of this connective structure. Here zinc in extracellular enzymes is used in cutting. Hence by a scissors (Zn) and paste (Cu) operations connective tissue could be built, destroyed and rebuilt to allow growth. Multicellular life could only arise when these two metals became sufficiently available. Copper also became valuable in the production of new messengers created by oxidation of peptides to give amidated units and in protective devices such as superoxide dismutase, see below.

Curiously it even seems reasonable to conclude that the concentrations of many free M^{2+} ions in the primitive sulfide sea were close to those in the primitive cytoplasm and they still are closely but not quite the same today [3] despite vast changes in the environment since the cytoplasm has not changed greatly in its reductive chemistry. What has had to change is the internal and external management of especially zinc and copper and the external capture of iron in vesicles and in extracellular fluids.

8.3. An unusual Cu/Zn enzyme

An exceptional enzyme in the cytoplasm which appeared late in evolution is the Cu/Zn superoxide dismutase (SOD). This protein replaced the Fe or Mn SOD of prokaryotes and we must ask why was this advisable. The Cu/Zn SOD is anomalous in that it is not just a particular type of coordination compound of Cu which is unique, especially so in the cytoplasm, using four histidine ligands, but it is also a unique coordination compound of zinc in the cytoplasm, also using four histidines and no thiolates. The only clear-cut difference in value is that the Mn and Fe SOD proteins tend to release free metal while Cu/Zn SOD is formed irreversibly and the metal ions are extremely strongly held. The Mn and Fe SOD proteins are produced in response to Mn or Fe deficiency due to the Fur transcription factor, suggesting that in prokaryotes Mn and/or Fe compounds are the major source of superoxide (from O_2). The Cu/Zn SOD synthesis depends on a specific supply of copper (see Section 5.5) so that maybe copper oxidases are the major risk for the production of superoxide in higher eukaryotes. We do know that much if not all copper is placed in enzymes by selective carriers, Fig. 31. The implication is that few if any copper enzymes are in equilibrium exchange. Maybe the only equilibration in one phase or compartment is between the buffer (say metallo-thionein or albumin), copper transcription factors and the two faces of copper pumps, respectively.

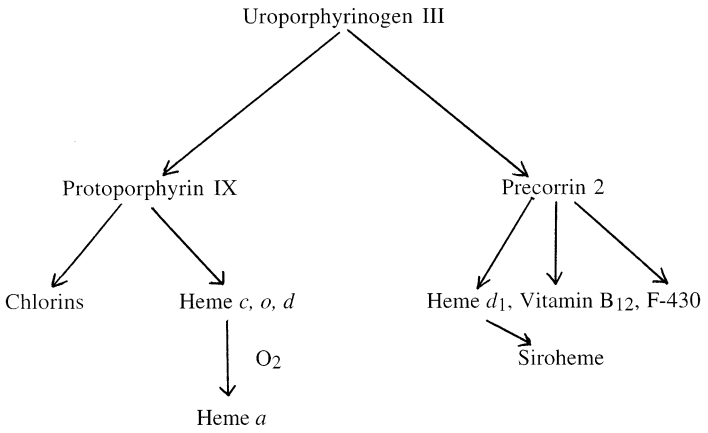


Fig. 34. The evolutionary relationship of the various Co, Ni, Fe and Mg ring complexes related to porphyrin.

8.4. The origin of porphyrin and its derivatives

The diversity of metal ring chelates derived from porphyrin are shown in Fig. 34 [77]. It is considered today that the cobalt cobalamins (B_{12}) were synthesised before the ring systems of iron (heme), nickel (F-430) or magnesium (chlorophyll) co-factors. In fact there are even today archae bacteria which do not synthesise heme iron but need coenzyme B_{12} compounds since DNA synthesis from RNA requires cobalamin in these primitive organisms [77].

The separation of the cobalt, iron, nickel and magnesium pathways is also shown in Fig. 34. The suggestion is that protoporphyrin IX is a relatively new pathway compared to that using precorrin-2. It is noteworthy that prokaryotes alone make cobalamin, F-430, and heme d_1, a siroheme. Siroheme handles sulfite and nitrite reduction and could have been valuable very early in evolution and its use certainly predated high levels of dioxygen in the atmosphere. No eukaryote synthesises these co-enzymes and many do not even require any of them as vitamins.

There are known to be two separate routes to vitamin B_{12} today, one anaerobic the other aerobic [77,78]. The ring closure step in both cases involves the loss of one methene bridge of the porphyrin ring as acetaldehyde in the case of the anaerobic path, the earlier path, and as acetic acid (with an absolute dependence on dioxygen) in the later aerobic path. Both pathways require the initial condensation of eight molecules of δ-amino laevulinic acid, as do all the four porphyrin cofactor syntheses, to give uroporphyrinogen III. This initial condensation reaction today requires a zinc enzyme. The further methylation reactions occur before the ring closure which in the aerobic pathway utilises an iron oxygenase. Thus the pathways have been adjusted in evolution to make use of the evolved environment. A good example is the production of heme *a* which requires dioxygen. Clearly this cannot be the result of random mutations [80].

A further distinction between the syntheses of these co-enzymes lies in the insertion steps for the metal elements [75,77]. The anaerobic pathway of insertion of cobalt does not require ATP and the insertion step is parallel to one of the iron insertion steps to form heme. The aerobic pathway requires ATP for insertion of cobalt and this route is parallel to that for insertion of magnesium in chlorophyll. A second enzyme which can act in the insertion of cobalt in anaerobic organisms is also capable of inserting iron into siroheme. Once again the pathways have evolved with the change of the environment.

The possible picture of evolution is therefore in the sequence (a), (b), (c) and then (d)

		Heme d_1				
Ring Chelate	Cobalamin,	Siroheme, (a)	F-430 →	Chlorophyll (b)	→ Heme(o,b_1d_1) (c)	⟶ Heme a (d)
Metal Ion	Co	Fe	Ni	Mg	Fe (Cu)	Fe.Cu
	RNA→DNA	SO_3^{2-}	H_2	Light	$O_2 \rightarrow H_2O$	$O_2 \rightarrow H_2O$
Substrates				Mn	NO(-Cu)	
	Methylation	NO_2^-		$H_2O \rightarrow O_2$		

The selection steps separating the different metal ions remain somewhat mysterious although there are now known three crystal structures of chelatases [75]. In the cobalt chelatase two histidines replace one histidine in the iron chelatase. The known chelatases split as follows

ATP-independent	Co(anaerobic), Fe	$(His)_{1\text{-}2}Glu$ binding site
ATP-dependent	Mg, Co (aerobic)	

It is not necessary that the insertion enzymes equilibrate with free metal ions but those for magnesium and iron do so. The Mg^{2+} insertion must use O-donors centres while the structure of the Fe^{2+}-chelatase shows it to be based on N/O-donation. Here the risk in the cell lies in the competition of the iron site with zinc since the iron chelatase can insert Zn^{2+}. The low binding constants for iron (and even zinc) of around 10^6 to 10^8 in the chelatase makes it possible to take up Fe^{2+} at a concentration level of about 10^{-7} M while the sequestration of Zn^{2+} by N,S-donors present in cells with binding constants $> 10^{11}$ makes Zn^{2+} unavailable (see Sections 4 and 5). The use of effective binding constants then follows the discussion given in the earlier section on the uptake of Fe^{2+} into Fe_3S_4 clusters and the binding differences between Fe^{2+} in Fur and Zn^{2+} in Zur.

A feature which makes the story of the evolution of these systems extremely puzzling is the large number of genes involved. There are at least fifteen genes for the synthesis of vitamin B_{12}, called *cob* genes [77]. Again the origin of their pathway is hard to explain in terms of random mutations [80].

8.5. Evolution of molybdenum, tungsten and selenium functions

While molybdenum has become a centrally important element in most organisms, tungsten, which shared a similar role with molybdenum in certain primitive organisms is not found to be of value in advanced living systems. Even so molybdenum has changed its role from being most frequently involved in reduction of N_2 and R–COOH, to reduction of SO_4^{2-} and NO_3^- and then to oxidation of R–CHO. A similar switch toward oxidative chemistry is to be seen in the uses of selenium. At first selenium was undoubtedly used in hydride transfer reactions acting as a more powerful catalysts than sulfur. With the advent of dioxygen the role of selenium has changed to that of a catalyst for the removal of peroxides and of iodine. Once again the special switches in chemistry of several elements with the coming of dioxygen suggest that evolution does not depend upon random mutation but that it operates following stress due to the environmental changes organisms suffer, see Section 9.1.

8.6. Na^+/K^+ and Ca^{2+} distribution and evolution

The fact that Ca^{2+} was rejected, of necessity, by early life forms allowed the later use of Ca^{2+} in assisting the development of eukaryotes, multi-compartmental single cell organisms, from prokaryotes. The calcium ion became the major second messenger in cells, as described above. The further fact that Na^+ was rejected, of necessity, by early life together with the acceptance of K^+ internally assisted the development late in evolution of nervous tissue and then of the brain. Of course this evolutionary progression was cooperative in the sense that the nerve Na^+/K^+ primary message interacted with the secondary Ca^{2+} message at nerve and nerve/muscle junctions, synapses. The rejection of two common cations Na^+ and Ca^{2+} and also one anion, chloride, gave evolution sources of stored electrolytic-chemical energy which became a major feature of the evolution of animals (Table 31). The use of the Na^+/K^+ gradient depended on the simultaneous evolution of a new

Table 31
Evolution of simple ionic equilibrium signals

Stage 1.	Primitive Organisms Prokaryotes	Mg^{2+}/ATP^{4-} controls phosphorylations Fe^{2+} controls redox equilibrium $Na^+/K^+/Cl^-$ control osmotic pressure
Stage 2.	Single cell Eukaryotes	Ca^{2+} controls activated states and relationship to environment Mn^{2+} controls development of plant-related organisms
Stage 3.	Multicellular	Zn^{2+} controls hormonal responses relating to growth of organisms and development. $Cu^+(Cu^{2+})$ and Zn^{2+} control connective tissue syntheses and degradations. Extended use of Ca^{2+} in excited state signals. Generation of Na^+ (K^+) signalling and the evolution of the nervous system.

Table 32
Maintained pathways throughout evolution

Pathway	Example
Syntheses and degradation of saccharides	Glycolysis (Mg^{2+})
Decarboxylic acid cycle	CO_2 incorporation later reversed to energy capture, Krebs cycle (Fe, Mg)
Amino acid synthesis	Products of glycolysis and Krebs cycle
Protein synthesis	Methionine initiation (Co, B_{12})
DNA, RNA syntheses	Nucleic acid pathways products of amino acids
Fat synthesis	β-carbon oxidation/reduction (flavin, Fe)
Nitrogen incorporation	Formation of NH_3 (Mg, Fe, Mo)
Hydrogen reactions	H_2 as a reductant (Fe, Ni)

pumping system the Na^+/K^+ ATP-ase. In fact mutation of the Na^+/K^+ ATP-ase leads to impairment of the neural plate during embryonic development and a very defective nervous system in the adult, if the organism survives [79].

9. Selection of elements: summary

We see that under given conditions of availability at a very early evolutionary stage, selection of elements for given purposes was achieved in the cytoplasm either through:

1. Thermodynamic competition with a limited supply of ligands, protein donors.
2. A kinetic pathway of transformations and transfers to permanent sites.

This selection in the cytoplasm, once made and codified, has been largely maintained throughout all subsequent stages of evolution even when availability changed dramatically (Table 32) since the demand is for a fixed set of reduction reactions leading to a fixed set of polymer types, DNA, RNA, proteins, fats, polysaccharides which remains as the basic necessity of life. The implication is that evolution largely depends upon adding-on, using environmental change, as opposed to making changes in existing cytoplasmic pathways. It is a form of organisational re-engineering. There are exceptional features of course but it is obvious that the major pathways are indeed highly conserved (Table 32). An analogy is that of the growth of a city which expands in the suburbs but keeps the centre largely the same. This means that the incorporation of metal ions into a large group of cytoplasmic centres which act as catalysts or controls has not been allowed to alter greatly even though the environment changed dramatically. There is then a necessity that despite the external environmental variations the cytoplasms of all cells have closely the same free metal ion concentrations but improved pathways of incorporation. Added-on systems in the cytoplasm therefore must either be quite new, non-interfering, kinetic pathways, (2) above, or if equilibrated pathways are used the binding constants of new metal ion systems in equilibrium, (1) above, must

be very similar to those already present. This is the essence of maintained cellular homeostasis through evolution when different elements became more or less available. To supply the cytoplasm it was also necessary not just to adjust uptake but also to devise export modes for essential elements to protect them from one another and to prevent newly introduced elements in the environment from entering the cytoplasm except in non-exchanging sequestered sites since they are bound to compete for sites with other elements.

Meanwhile evolution develop a great range of added-on systems based upon new reactions mainly outside the cytoplasm with new connections across new membranes. Here it is the development of compartments not linked directly to the cytoplasm and not interfering with its reaction pathways that are dominant. The compartments could be captured organelles, vesicles or extracellular fluids. In organelles the redox chemistry and uses of metal ions is in particular cases very different from that in the cytoplasm. Noticeably we see this in the extracellular biological chemistry of copper and to some degree zinc which became major elements in vesicles and outside cells connected to the synthesis and degradation of connective tissue and to signalling between cells (Table 33).

While early in evolution the elements described above were difficult to obtain, e.g. Zn^{2+}, and others impossibly insoluble, e.g. Cu, yet others were too concentrated in the sea. In particular Na^{+}, Ca^{2+} and Cl^{-} had to be pumped out of cells so that the osmotic pressure and the solubility of organic molecules could be protected. This has meant that from the most primitive cells to the most advanced the cytoplasm of the cell has been kept at 10^{-2} M Na^{+}, 10^{-2} M Cl^{-} and well below 10^{-5} M Ca^{2+}, while K^{+} is at 10^{-1} M and Mg^{2+} is 10^{-3} M. The consequences of the creation of gradients of these cations have been profound for message transmission and also for mineralisation.

Table 33
Examples of final distributions of metals and ligands

(a) *Inside cells*	
$Zn^{2+}(10^{-11}$ M)	$Cu^{+}(10^{-15}$M)
$(RS^{-})_4$ and N/RS^{-}	$(RS^{-})_4$ $(RS^{-})_2$
4/5 coordinate	3 or 4 coordinate.
3/4 His (input channels)	$(imid^{-})(imid)_3$ SOD
	4 coordinate distorted tetragonal
Exchange intermediate	Exchange slow
(b) *Outside cells*	
$Zn^{2+}(<10^{-11}$ M)	$Cu^{2+}(<10^{-15}$M)
$(RS^{-})_2$(imid) pump	$(RS^{-})_2$ pump
Albumin	Albumin N^{-}/N
	$RS^{-}(imid)_2$
(c) *In vesicles*	
Zn^{2+} (10^{-3} to 10^{-11} M)	$Cu(10^{-15}$ M)
free or loosely bound	$(RS^{-})_2$ Met
3(imid) (H_2O)	

A major question which arises is how was it possible in evolution to maintain much of the chemistry of the cytoplasm while that of vesicles and extracellular fluids developed dramatically. A possible explanation is that environmental changes induced damage but only to certain vulnerable proteins. The response of the code is of course to replace the damaged proteins requiring the DNA to cycle more rapidly through double then single stranded conformations. Single stranded DNA is itself more vulnerable to damage and hence mutation and other factors, see next section, caused proteins to evolve to counter stress through localised changes of DNA [80]. The major stresses imposed on organisms are caused by new reagents such as O_2, NO and Zn^{2+} and Cu^{2+}. Hence these elements (compounds) have a major influence on the evolution of biological species.

A final word is necessary concerning the selective distribution of all the elements. Much of this article has concentrated on the static picture of the steady homeostatic state. The dynamic model of changing steady states is the true one in which the distributions are maintained by different flow rates in and out of compartments and binding sites. The elements for which there are both pumps and channels flow across membranes in non-random patterns while those which flow in the cytoplasm only are virtually in random motion. Finally there are elements for which flow is unimportant. At the present time this topic is little studied but it would appear that flow patterns across membranes are most dominant for Na^+, K^+, H^+, and Ca^{2+}. Flow in the cell cytoplasm is strongest for K^+, Na^+, H^+, Mg^{2+}, Mn^{2+} and Fe^{2+} and slower flow for Zn^{2+} which is of considerable importance. The most static elements are Cu^+, Ni^{2+}, Co^{2+} and Mo. The non-metals flow in mobile co-enzymes. Lastly the electron flows in many redox proteins especially in membranes.

9.1. Note: modern views as to the nature of adaptable DNA

The multitude of ways in which the coded information in a cell may evolve is described and analysed in Ref. [80]. It is surprising that the DNA code appears to be sensitive not just to random mutation but to stress-induced change including mutation and transpositional alteration which have different rates under special states of a cell cycle. There is also a consensus view in the volume that a genome encodes efficient strategies for evolution, in other words that there is some kind of natural genetic engineering involving not just localised base changes but recombination, transposition and duplication of large pieces of changing sequence. The outcome of this flexibility of the code in the face of inorganic external environmental circumstances is not yet open to full analysis but we believe it is the basis of evolution as explained above.

Acknowledgements

We acknowledge the help of Dr J. Armando, Superior Técnico, Lisboa, for his helpful assistance with the data collected in Section 5.4. We thank Professor Sir Phillip Cohen, University of Dundee, for his help with Section 5.15.

References

[1] J.J.R. Fraústo da Silva and R.J.P. Williams, The Biological Chemistry of the Elements, Oxford University Press, Oxford, 1991. Reprinted with additions, 1996.
[2] R.J.P. Williams, J.J.R. Fraústo da Silva, The Natural Selection of the Chemical Elements, Oxford University Press, Oxford, 1996.
[3] R.J.P. Williams, J.J.R. Fraústo, Bringing Chemistry to Life: From Matter to Man, Oxford University Press, Oxford, 1999.
[4] D. Latchman, Gene Regulation, 3rd ed., Stanley Thornes, Cheltenham, UK, 1998.
[5] B. Lewin, Genes, 6th ed., Oxford University Press, Oxford, 1996.
[6] G.L. Zubay, Biochemistry, 4th ed., Brown, Dubuque IA, USA, 1998.
[7] L. Stryer, Biochemistry, 3rd ed., Freeman, New York, 1988.
[8] F.C. Neidhardt, R. Curtiss, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, H.E. Umbarger (Eds.), *Escherichia coli* and Salmonella, 2nd ed., American Society of Microbiology Press, Washington, USA, 1996.
[9] A. Cornish-Bowden, Fundamentals of Enzyme Kinetics, Portland Press, London, 1995.
[10] C. Dennison, A. Berg, G.W. Canters, 36 (1997) 3262 and references therein.
[11] B. Schwaller, Presentation at the 5th European Meeting on Calcium Binding Proteins, Munster, Germany, 1998.
[12] B. Magasanik, in: F.C. Neidhardt, R. Curtiss, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, H.E. Umbarger (Eds.), *Escherichia coli* and Salmonella, 2nd ed., American Society of Microbiology Press, Washington, USA, 1996, pp. 1344–1356.
[13] G.A. Marzluf, Ann. Rev. Microbiol. 57 (1997) 73.
[14] D.M. Lawson, C.E.M. Williams, L.A. Mitchenall, R.N. Pau, Structure 6 (1998) 1529 in press.
[15] R.B. Gennis, V. Stewart, in: F.C. Neidhardt, R. Curtiss, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, H.E. Umbarger (Eds.), *Escherichia coli* and Salmonella, 2nd ed., American Society of Microbiology Press, Washington, DC, 1996, pp. 217–261.
[16] P.C.H. Mitchell, R.J.P. Williams, J. Chem. Soc. (1962) 4570.
[17] K.V. Rajagopalan, in: F.C. Neidhardt, R. Curtiss, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, H.E. Umbarger (Eds.), *Escherichia coli* and Salmonella, 2nd ed., American Society of Microbiology Press, Washington, DC, 1996, pp. 674–679.
[18] (a) D.A. Draper, V.K. Misra, Nat. Struct. Biol. 5 (1998) 927. (b) C.S.G. Phillips, R.J.P. Williams, Inorganic Chemistry, vol. II, Oxford University Press, Oxford, 1966, p. 80–90.
[19] R.J.P. Williams, J. Theoret. Biol. 1 (1961) 1.
[20] N.J. Birch (Ed.), Magnesium and the Cell, Academic Press, London, 1993.
[21] E. Carafoli, C. Klee (Eds.), Calcium as a Cellular Regulator, Oxford University Press, Oxford, 1999.
[22] S. Wang, W.R. Trumble, H. Liao, C.R. Wesson, A.K. Dunker, C. Kung, Nat. Struct. Biol. 5 (1998) 470.
[23] L.G. Sillen, A.E. Martell (Eds.), Stability Constants, Special Publication No. 17, The Chemical Society, London, 1964.
[24] J. Zhou, Z. Hu, E. Münck, R.H. Holm, J. Am. Chem. Soc. 118 (1996) 1966.
[25] M. Eigen, Z. Electrochem. 64 (1960) 115.
[26] R.J.P. Williams, Biol. Rev. 28 (1953) 381.
[27] S. Silver, in: F.C. Neidhardt, R. Curtiss, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, H.E. Umbarger (Eds.), *Escherichia coli* and Salmonella, 2nd ed., American Society of Microbiology Press, Washington, DC, 1996, pp. 1091–1102.
[28] R.J.P. Williams, FEBS Lett. 140 (1982) 3.
[29] T.V. O'Halloran, Science 261 (1993) 715.
[30] R.J.P. Williams, Biochim. Biophys. Acta 1448 (1998) 153.
[31] T. Saito, M.R. Wormald, R.J.P. Williams, Eur. J. Biochem. 197 (1991) 29.
[32] A. Adrait, L. Jacquamet, L.L. Pape, A. de Peredo, D. Aberdam, J.-L. Hazemann, J.M. Latour, I. Michaud, Biochemistry 38 (1999) 6248.

[33] X. Qiu, C.L.M.J. Verlinde, S. Zhang, M.P. Schmitt, R.K. Holmes, W.G.T. Hol, Structure 3 (1995) 87.
[34] H. Beinert, M.C. Kennedy, C.D. Stout, Chem. Rev. 96 (1996) 2335.
[35] J. Green, J.R. Guest, FEBS Lett. 329 (1993) 55.
[36] H. Komeda, M. Kobayashi, S. Shimizu, Proc. Natl. Acad. Sci. USA 94 (1997) 36.
[37] M. Kobayashi, S. Shimizu, Eur. J. Biochem. 261 (1999) 1.
[38] E.R. Still, R.J.P. Williams, J. Inorg. Biochem. 13 (1980) 35.
[39] U. Kramer, J.D. Cotter-Howells, J.M. Charnock, A.J.M. Baker, J.A.C. Smith, Nature 379 (1996) 635.
[40] M.D. Harrison, C.E. Jones, C.T. Dameron, J. Biol. Inorg. Chem. 4 (1999) 145.
[41] A. Klug, J.W.R. Schwabe, Proc. Fed. Am. Soc. Expt. Biol. 9 (1995) 597.
[42] R.D. Palmiter, S.D. Findley, Proc. Natl. Acad. Sci. USA 93 (1996) 14934.
[43] D. Lu, B. Boyd, C. Lingwood, J. Biol. Chem. 272 (1997) 29033.
[44] C. Thelwell, N.J. Robinson, J.S. Turner-Cavet, Proc. Natl. Acad. Sci. USA 95 (1998) 10728.
[45] R.H. Holm, E.I. Solomon, Chem. Rev. 96 (1996) 2239.
[46] H.A.O. Hill, P.J. Sadler, A.J. Thomson (Eds.), Structure and Bonding, vol. 89, Springer, Berlin, 1997.
[47] C.B. Rensing, B. Mitra, B.P. Rosen, Proc. Natl. Acad. Sci. USA 94 (1997) 14326.
[48] (a) R.D. Graham, R.J. Hannam, N.C. Uren (Eds.), Manganese in Soils and Plants, Kluwer, Dordrecht, Netherlands, 1988. (b) L. Zhang, M.J. Crossley, N.E. Dixon, P.J. Ellis, M.L. Fisher, G.F. King, P.E. Lilley, D. MacLachlan, R.J. Pace, H.C. Freeman, J. Biol. Inorg. Chem. 3 (1998) 470.
[49] J.P. Griffith, J.L. Kim, E.E. Kim, M.D. Sintchak, J.A. Thomson, M.J. Fitzgibbon, M.A. Fleming, P.R. Caron, K. Hsiao, M.A. Navia, Cell 82 (1995) 507.
[50] R.A. Bradshaw, W.W. Brickey, K.E. Walker, TIBS 23 (1998) 263.
[51] E. Jabri, M.B. Carr, R.P. Hausinger, P.A. Karplus, Science 268 (1995) 998.
[52] (a) B.L. Vallee, D.S. Auld, Biochemistry 29 (1990) 5647. (b) B.L. Vallee, D.S. Auld, Biochemistry 32 (1993) 6493.
[53] S. Melino, C. Capo, B. Dragini, A. Aceto, R. Petruzzelli, TIBS 23 (1998) 381.
[54] C.P. Bobel, Cell 90 (1997) 589.
[55] N.M. Hooper, FEBS Lett. 354 (1994) 1.
[56] W. Stöcker, F.-X. Gomis-Rüth, W. Bode, R. Zwilling, Eur. J. Biochem. 214 (1993) 215.
[57] L. Aravind, E.V. Koonin, TIBS 23 (1998) 469.
[58] H.B. Gray, B.G. Malmström, R.J.P. Williams, J. Biol. Inorg. Chem. (2000) accepted for publication.
[59] R.J.P. Williams, Endeavour 8 (1984) 65.
[60] P.J. Keeling, W.F. Doolittle, TIBS 21 (1998) 285.
[61] P. Sadler (Ed.), Structure and Bonding, vol. 88, Springer, Berlin, 1997.
[62] H.A.O. Hill, P.J. Sadler and A.J. Thomson (Eds.), Structure and Bonding, vol. 90, Springer, Berlin, 1997.
[63] B.L. Vallee, R.J.P. Williams, Chem. Brit. 4 (1968) 397.
[64] R.J.P. Williams, FEBS Lett. 140 (1982) 3.
[65] F.S. Armstrong, R.J.P. Williams, FEBS Lett. 451 (1999) 91.
[66] J.E. Baldwin, C.J. Schofield, R.J.P. Williams, in preparation.
[67] P.L. Roach, I.J. Clifton, V. Fülop, K. Harlos, G.J. Barton, J. Hadju, I. Andersson, C.J. Schofield, J.E. Baldwin, Nature 375 (1995) 700.
[68] M.B. Robin, P. Day, Adv. Inorg. Chem. Radiochem. 10 (1967) 247.
[69] M. Kennedy, L. Yu, M.J. Lima, C.S. Ascenso, C. Czaja, I. Moura, J.J.G. Moura, F. Rusnak, J. Biol. Inorg. Chem. 3 (1998) 643.
[70] J.S. Valentine, E.B. Gralla, Science 278 (1997) 817.
[71] M.K. Chan, S. Mukund, A. Kletzin, M.W.W. Adams, D.C. Rees, Science 267 (1995) 1463.
[72] A. Volbeda, M-H Charon, C. Piras, E.C. Hatchikian, M. Frey, J.C. Fontecilla-Camps, Nature 373 (1995) 580.

[73] Chem. Revs. 96 (1996) 2315. A volume devoted to metallo-enzymes and includes much material relevant to the last two sections.
[74] S. Al-Karadagli, M. Hansson, B. Jönsson, L. Hederstedt, Structure 5 (1997) 1501.
[75] S.A. Chervitz, L. Aravind, G. Sherlock, C.A. Ball, E.V. Koonin, S.S. Dwight, M.A. Harris, K. Dolinski, T. Smith, S. Wong, J.M. Cherry, D. Botstein, Science 282 (1998) 2022.
[76] N.D. Clarke, J.M. Berg, Science 282 (1998) 2018.
[77] S.I. Beale, in: F.C. Neidhardt, R. Curtiss, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, H.E. Umbarger (Eds.), *Escherichia coli* and Salmonella, 2nd ed., American Society of Microbiology Press, Washington, DC, 1996, pp. 731–748.
[78] E. Raux, H.L. Schubert, J.M. Roper, K.S. Wilson, M.J. Warren, Bioinorg. Chem. 27 (1999) 753.
[79] C.S. Davies, N.J. Messenger, R. Craig, A.E. Warner, Dev. Biol. 174 (1996) 431.
[80] L.H. Caporale (Ed.), Molecular strategies in biological evolution, Ann. N.Y. Acad. Sci. 870 (1999).
[81] J.N. Butt, Ph.D. Thesis, University of California, Irvine, CA, 1993.
[82] H. Gunshin, B. MacKenzie, U.V. Berger, Y. Gunshin, M.F. Romero, W.F. Brown, S. Nussberger, J.A. Gollan, M.A. Heidiger, Nature 388 (1997) 482.