SPRINGER BRIEFS IN MOLECULAR SCIENCE
BIOMETALS

Luana Presta
Marco Fondi
Giovanni Emiliani
Renato Fani

Molybdenum Cofactors and Their role in the Evolution of Metabolic Pathways

Springer

SpringerBriefs in Molecular Science

Biometals

More information about this series at http://www.springer.com/series/10046

Luana Presta · Marco Fondi
Giovanni Emiliani · Renato Fani

Molybdenum Cofactors and Their role in the Evolution of Metabolic Pathways

Springer

Luana Presta
Laboratory of Microbial and Molecular Evolution, Department of Biology
University of Florence
Sesto Fiorentino
Italy

Marco Fondi
Laboratory of Microbial and Molecular Evolution, Department of Biology
University of Florence
Sesto Fiorentino
Italy

Giovanni Emiliani
Trees and Timber Institute National Research Council, IVALSA-CNR
Trento
Italy

Renato Fani
Laboratory of Microbial and Molecular Evolution, Department of Biology
University of Florence
Sesto Fiorentino
Italy

ISSN 2191-5407 ISSN 2191-5415 (electronic)
SpringerBriefs in Molecular Science
ISSN 2212-9901
SpringerBriefs in Biometals
ISBN 978-94-017-9971-3 ISBN 978-94-017-9972-0 (eBook)
DOI 10.1007/978-94-017-9972-0

Library of Congress Control Number: 2015938725

Springer Dordrecht Heidelberg New York London

Printed on acid-free paper

Springer Science+Business Media B.V. Dordrecht is part of Springer Science+Business Media (www.springer.com)

Contents

Chapter 1
Introduction

Abstract Molybdenum (Mo) is essential for (nearly) all biological systems, both Eukaryotes and Prokaryotes. Molybdenum, indeed, bounds to a pterin and forms the molybdenum cofactor (Mo-co) that, in different variants, represents the active compound at the catalytic site of all molybdenum-containing enzymes in nature. In some microorganisms, tungsten (W) is also coordinated by pyranopterin (W-co). The exchangeable use of these two elements probably traces its origin very far into the past.

Several transition elements are essential for microorganisms; among these, molybdenum (Mo) is essential for (nearly) all biological systems, both eukaryotes (including animals, plants, fungi), and prokaryotes (Archaea and Bacteria). From a biological viewpoint, Mo belongs to the group of trace elements; indeed it is required by organisms only in minute amounts. In spite of this, unavailability of Mo is lethal for microorganisms. Further, even when Mo is available in the medium it needs to be complexed by a special cofactor in order to confer catalytic activity to several enzymes. Molybdenum, in fact, is bound to a unique molecule, pterin, thus forming the molybdenum cofactor (Mo-co) (Fig. 1.1). Excluding bacterial and archaeal molybdenum nitrogenase, which is a multinuclear iron (Fe)–Mo-cofactor, all Mo-dependent enzymes (molybdoenzymes) hold a mononuclear form of the Mo-co, which consists of Mo coordinated to an organic tricyclic pyranopterin moiety, referred to as molybdopterin (Mendel and Bittner 2006; Schwarz 2005).

This cofactor, in different variants, represents the active compound at the catalytic site of all molybdenum-containing enzymes in nature (except the nitrogenase), that is several enzymes catalyzing diverse key oxotransfer reactions in the global carbon, sulphur, and nitrogen metabolism (Mendel and Bittner 2006). It serves as a scaffold to bind Mo and allows the metal to catalytically cycle through the IV, V, and VI oxidation states (Magalon et al. 2011). Importantly, it enables pterin-based enzymes to connect two electron Mo-co-catalyzed reactions to an additional active site via an electron transfer relay (ETR) able to sequentially transfer single electrons to the distal second site (Bertero et al. 2003, 2005).

L. Presta et al., *Molybdenum Cofactors and Their role in the Evolution of Metabolic Pathways*, SpringerBriefs in Biometals, DOI 10.1007/978-94-017-9972-0_1

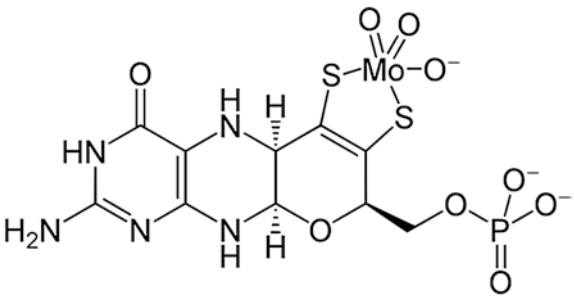

Fig. 1.1 Chemical structure of molybdenum-cofactor

In spite of the fact that it is only a minor constituent of the Earth's crust, molybdenum is the most abundant transition metal in present-day seawater (Anbar 2008) and, indeed, it is readily available (in an oxidizing environment) due to its presence as a trace element in aquatic environments in the anionic form, i.e. molybdate (MoO_4^{2-}).

In some microorganisms (mostly in thermophilic Archaea), tungsten (W) is also coordinated by pyranopterin (W-co) (Chan et al. 1995; Huber et al. 2009); indeed, prokaryotic species that do not require Mo use tungsten (as tungstate, WO_4^{2-}), which lies immediately below molybdenum in the periodic table (Hille 2002). In addition, W can be selectively transported into prokaryotic cells by some transporters (Bevers et al. 2009; Makdessi et al. 2004; Andreesen and Makdessi 2008; Grunden et al. 1996) and is an essential element for enzymes belonging to the aldehyde:ferredoxin oxidoreductase (AOR) family (Makdessi et al. 2004; Andreesen and Makdessi 2008). Since Mo and W share chemical and physical similarities, it is often almost impossible to distinguish the utilization of these two elements based on the sole analysis of proteins sequence and/or metabolic pathways. The exchangeable use of these two elements probably traces its origin very far into the past. This relies on the idea that life probably arose in anaerobic conditions where tungsten might have been the first of the two elements to be acquired by living organisms (Magalon et al. 2011). This intriguing theory will be further discussed in the next chapter; in order to better understand how molybdenum-based systems arose we will focus on geochemical aspects of this element.

Bibliography

Anbar, A. D. (2008). Elements and evolution. *Science, 322*(5907), 1481–1483.

Andreesen, J. R., & Makdessi, K. (2008). Tungsten, the surprisingly positively acting heavy metal element for prokaryotes. *Annals of the New York Academy of Sciences, 1125*(1), 215–229.

Bertero, M. G., Rothery, R. A., Palak, M., Hou, C., Lim, D., Blasco, F., et al. (2003). Insights into the respiratory electron transfer pathway from the structure of nitrate reductase A. *Nature Structural & Molecular Biology, 10*(9), 681–687.

Bertero, M. G., Rothery, R. A., Boroumand, N., Palak, M., Blasco, F., Ginet, N., et al. (2005). Structural and biochemical characterization of a quinol binding site of *Escherichia coli* nitrate reductase A. *Journal of Biological Chemistry, 280*(15), 14836–14843.

Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W., & Rees, D. C. (1995). Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science, 267*(5203), 1463–1469.

Huber, R., Sch Bevers, L. E., Hagedoorn, P. L., & Hagen, W. R. (2009). The bioinorganic chemistry of tungsten. *Coordination Chemistry Reviews, 253*(3), 269–290.

Magalon, A., Fedor, J. G., Walburger, A., & Weiner, J. H. (2011). Molybdenum enzymes in bacteria and their maturation. *Coordination Chemistry Reviews, 255*(9), 1159–1178.

Makdessi, K., Fritsche, K., Pich, A., & Andreesen, J. R. (2004). Identification and characterization of the cytoplasmic tungstate/molybdate-binding protein (Mop) from *Eubacterium acidaminophilum*. *Archives of Microbiology, 181*(1), 45–51.

Mendel, R. R., & Bittner, F. (2006). Cell biology of molybdenum. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1763*(7), 621–635.

Schwarz, G. (2005). Molybdenum cofactor biosynthesis and deficiency. *Cellular and Molecular Life Sciences CMLS, 62*(23), 2792–2810.

Chapter 2
Molybdenum Availability in the Ecosystems (Geochemistry Aspects, When and How Did It Appear?)

Abstract The distribution of life on Earth is constrained by the distribution of 20 bio-essential nutrients. Particularly, the availability of those elements required as cofactor might have influenced the timing of origin and evolution of metabolic pathways. Since the amount of bio-available Mo strongly changed during evolution to form anoxic to oxygenic environments, it is interesting to speculate about the appearance of Mo containing enzymes in relation to other metals with similar chemical characteristics.

2.1 Elements and Evolution of Biological Systems

The current ecosystem is the result of the Earth 4.5 billion years of evolutionary history and has been modelled thanks to the combined effects of tectonic, photochemical and biological processes. The abundance of the elements in the universe and the way in which the planetary system was formed controlled the abundance of metals on the primitive Earth, which was (and still is) a metal rich planet (Prantzos et al. 1994). The fact that the metal elements and hydrogen were slightly in excess in respect to non-metals (such as oxygen) has very likely determined the initial reducing capacity of condensed matter.

Among the most important and relatively abundant non-metallic elements, oxygen appeared in liquid water and in solid (mainly in low-valent metal oxides or mixed oxides such as silicates); nitrogen in the atmosphere formed as N_2 or NH_3; carbon as CO, CO_2 (a large part of the carbon has always been oxidized to CO_2) and CH_4, chlorine as HCl initially (if out of contact with metal oxides), and sulfur as metal sulfides and H_2S. Much of the carbon could not be reduced and remained as CO or CO_2. Oxidation of sulphur and nitrogen was prevented by the overall lowered availability of molecular oxygen due to its high affinity for metals, hydrogen and carbon. Metals, which could not be reduced and were kept in their highest oxidation states occurred in large part as simple ions in oxide lattices or soluble salts. This is the case for Al^{3+}, Ca^{2+}, Mg^{2+}, Na^{+} and K^{+}. Some of them readily gave

L. Presta et al., *Molybdenum Cofactors and Their role in the Evolution of Metabolic Pathways*, SpringerBriefs in Biometals, DOI 10.1007/978-94-017-9972-0_2

insoluble mixed oxide based salts such as silicates and carbonates. The most soluble oxides then reacted with the limited amounts (by abundance) of HCl to give soluble chlorides (such as NaCl and KCl) in the sea once water condensed. Since oxygen was not in sufficient abundance those metals with a relatively low affinity for oxygen could not form oxides and they formed sulphides (e.g. of many transition metals) or metals and alloys (e.g. metals such as gold). Initially, there would have been no excess of oxygen or sulphur as elements though both were present as hydrides, H_2O and H_2S. All the associations in the exposed surface zones are firmly based on the restrictions of variables due to total composition (abundances) and certain equilibria at room temperature and pressure.

These concentrations of all important elements through time (shown in Fig. 2.1) along with their oxidation states represented extremely strong constraints upon the formation of an energized system. Thus, it was the combination of both of them that eventually led to the formation of organic systems which predated life or which grew simultaneously as encoded life.

2.1.1 Early Atmosphere and Early Sea Composition

Even though there is not a complete agreement on the composition of the **early atmosphere** it is likely that it was somewhat reducing since it contained H_2, CH_4, CO and H_2S in addition to CO_2 (Kasting 1993). On the basis of all the available data it is still not possible to state whether CO and CO_2 (oxidizing gases) were or were not in thermodynamic equilibrium with the reducing gases H_2 and CH_4 and with H_2O, given the uncertainty in their relative abundances. Also, the nitrogen content is uncertain since there could have been such gases as N_2, NH_3 and HCN. The exact partial pressure of these gases is unknown but it is possible that the most represented were N_2 and CO_2 with considerable amounts of CO. For clarity a table with the major constituents of early atmosphere is reported in Table 2.1.

Availability of different elements in the **early sea** is much more difficult to formulate since it depended on the solubility of the different compounds. The vast majority of abundant metal ions form rather insoluble salts with abundant anions such as silicate (Mg^{2+}, Ca^{2+}, Al^{3+}) or sulphide (transition metal ions). The remaining few abundant metal ions are Na^+ and K^+, and one anion, Cl^-, which do not easily give precipitates with abundant anions or cations respectively. As a consequence, especially Na^+ and K^+, and Cl^- were left in solution in the sea and were limited there by the lower abundance of chlorine relative to the combined abundances of Na and K. Almost all chlorine on Earth is still in the seas and oceans or it evaporates today as NaCl. A small quantity of bicarbonate/carbonate accompanied by the metal elements, Mg^{2+} and Ca^{2+}, which form somewhat soluble salts, was also present. According to the idea of a reducing atmosphere it is very likely that no sulphate or nitrate were present since they were too oxidizing. Anyway, the transition metal elements, both silicates and sulphides are insoluble; nevertheless with increasing atomic number in the

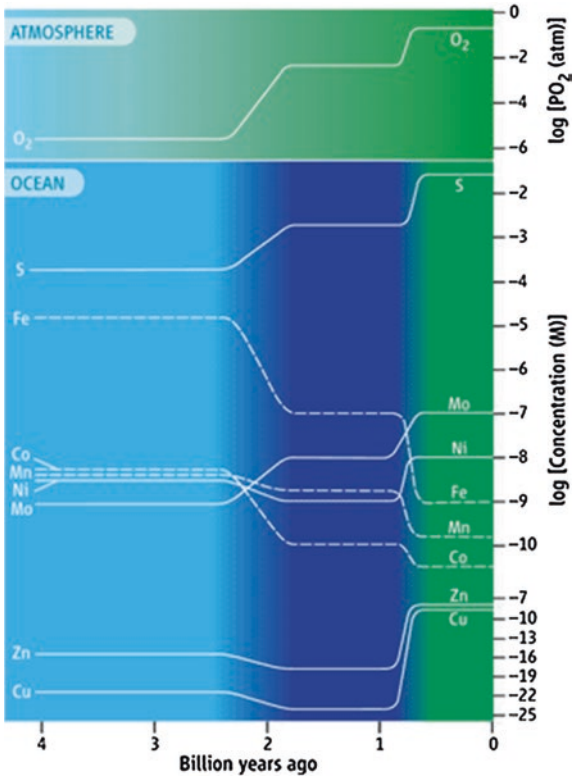

Fig. 2.1 Changes in element abundances through time. These histories are approximate, based on simple geochemical models and inferences from ancient sediments. An expansion in H_2S-rich ocean regions after 2.4 billion years ago is assumed. Color gradations indicate a transition from anoxic, S-poor oceans before 2.4 billion years ago (*light blue*) to H_2S-rich oceans between 1.8 billion and 800 million years ago (*dark blue*), subsequently giving way to complete ocean oxygenation (*green*). Different line styles are for clarity only; *dashed lines* are for elements with falling concentrations. Reproduced from Anbar (2008). With kind permission of © Science. CREDIT: Susan Selkirk/Arizona State University

Table 2.1 Principal constituents of early atmosphere

Constituents	Percent by volume
CO_2	91
N_2	6.4
H_2S	2.0
HCI	0.2
H_2O	0.2
NH_3	0.01
CH_4	0.02
CI_2	0.01

Data are taken from Rubey (1955)

transition metal series from Mn to Cu the sulphides become the dominant species. The insolubility of precipitated sulphides was responsible for the absence of copper and initially of molybdenum and the very low content of zinc, on the contrary iron as Fe^{2+}, Mn^{2+} and even tungsten (W) remained in soluble forms in the primitive sea. The approximate probable concentrations of all of the major ions in the primitive sea under either anaerobic or aerobic conditions is shown in Table 2.2.

However, the composition of both the early sea and atmosphere might have changed since the different compounds were not in temperature equilibrium with the hot core and the very high temperature on Earth's surface. As a consequence, the environment was liable to thermal and photochemical kinetic changes even before the appearance of the first living forms.

The gases H_2 and O_2 are more prone to react with other chemicals than with themselves, so that the reverse reaction did not appear before the changing in the oxidation state of the earth. The kinetic and sustained traps for hydrogen were mainly light non-metals such as carbon and nitrogen while the traps for oxygen were heavier non-metals such as sulphur and then various metal sulphides such as V, Cr, Mn, Fe, Co, Ni, Cu, Zn and Mo slowly producing more soluble oxides. A feature of the states of these metals is that over a reasonable period of time they equilibrate with oxygen and amongst themselves but not with C/H compounds so that we can believe that the sea and atmosphere changed continuously through an increase in redox potential equilibrated with the slowly increasing O_2 partial pressure. At the same time more reducing equivalents were trapped (stored) in organic molecules and the carbon and nitrogen compounds in the atmosphere were also oxidised so that the atmosphere became increasingly composed by CO_2 and N_2. Later still surface sulphides were oxidized to sulphate and N_2 to some nitrate in the sea.

Table 2.2 Approximate probable concentrations of the major ions in aerobic and anaerobic conditions in the primitive seas

Metal ion	Anaerobic conditions (M)	Aerobic conditions (M)
Na^+	$>10^{-1}$	$>10^{-1}$
K^+	10^{-2}	10^{-2}
Mg^{2+}	10^{-2}	10^{-2}
Ca^{2+}	10^{-3}	10^{-3}
Mn^{2+}	10^{-6}	10^{-6}
Fe	10^{-7} (Fe^{II})	10^{-19} (Fe^{III})
Co^{2+}	$<10^{-9}$	10-set
Ni^{2+}	$<10^{-9}$	$<10^{-9}$
Cu	$<10^{-20}$, Cu^{I}	$<10^{-10}$, Cu^{II}
Zn^{2+}	$<10^{-12}$	$<10^{-8}$
Mo	**$<10^{-10}$ (MoS_4^{2-})**	**10^{-8} (MoO_4^{2-})**
W	**$<10^{-9}$ (WS_4^{2-})**	**10^{-9} (WO_4^{2-})**
H_2S	10^{-2}	$<10^{-10}$
HPO_4^{2-}	$<10^{-3}$	$<10^{-3}$

Moreover, the distribution of bio-essential elements in early sea, is not easy to be studied, but inferences can be drawn from the characteristics of rocks formed from seafloor sediments. The best example is the abundance of iron in the geologic record (Isley and Abbott 1999), massive deposits of iron sedimentary ore older than 1.8 billion years. Such deposits almost disappear from the more recent record. The analysis of Banded Iron Formations (BIFS) suggests that the iron-rich oceans in the first half of Earth's history gave way to later iron scarcity (Poulton et al. 2004; Holland 2006). The changes in the availability of other bio-essential elements are subtler and must be inferred by sophisticated methods of analysis.

Due to recent instrumental analytical advances it was possible to identify new paleoredox state markers (e.g. molybdenum isotopes, Re, Os), and to improve existing ones (e.g. Mo-TOC, Re/Mo). In this sense, the sedimentary molybdenum is commonly used as a paleoredox conditions marker. An enrichment of this element in sediments and sedimentary rocks, is interpreted as reflecting low potentials redox in the environment (Sageman and Lyons 2003; Tribovillard et al. 2006; Pearce et al. 2008; Böning et al. 2008). Thus, anoxic facies are generally enriched in molybdenum with respect to oxic and suboxic ones and the mechanics of this enrichment process have been analysed extensively (e.g. Emerson and Huested 1991; Crusius et al. 1996; Helz et al. 1996; Zheng et al. 2000). The mobilization of molybdenum enrichment in the marine sedimentary environment also produces an isotope fractionation of mercury (Siebert et al. 2003; Arnold et al. 2004; Poulson et al. 2006; Pearce et al. 2008; Neubert et al. 2008; Archer and Vance 2008). Thus, changes in redox conditions in the marine depositional environment can be determined by the isotopic composition of molybdenum.

2.1.2 Elements Avaibility and Emergence of Life

Since the onset of primordial living systems, which most likely occurred around 3.8–3.5 billion years ago (Lazcano and Miller 1996), nearly all elements have been used by microbial metabolic activities. From a geochemical perspective the cycles of energy and minerals (especially H, C, N, O and S) required by vital processes and the biogeochemistry processes can be considered thermodynamically constrained redox reactions catalyzed by microbial metabolic pathways (Falkowski et al. 2008). Thus, the biogeochemical cycles might be interpreted as the interconnection between abiotic driven acid-base reactions (acting on the geological time scale to supply the system with elements through volcanism and rock weathering) and biotic redox reactions. Such metabolic reactions transformed the redox state of the planet into an oxidative environment once oxygenic photosynthesis appeared in cyanobacteria about 3–2.7 billion years ago (Canfield 2005). The appearance of this crucial metabolic ability led to a progressively strong increase of the O_2 concentration in the atmosphere in a relatively short timescale between 2.4 and 2.2 billion years ago, as shown in Fig. 2.2 (Bekker et al. 2004). Depending on the composition and original redox state of the prebiotic atmosphere

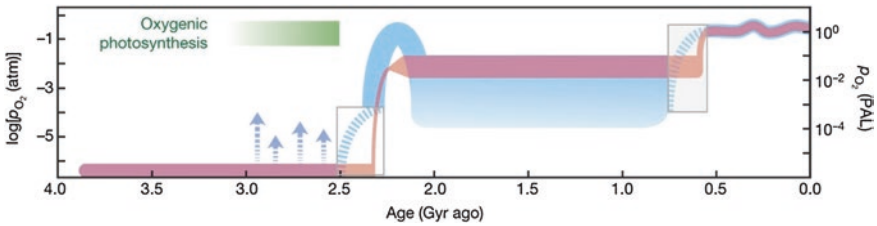

Fig. 2.2 Evolution of Earth's atmospheric oxygen content through time. The faded red curve shows a 'classical, two-step' view of atmospheric evolution, while the *blue curve* shows the emerging model (PO_2, atmospheric partial pressure of O_2). *Right axis* PO_2 relative to the present atmospheric level (PAL); *left axis* log PO_2. *Arrows* denote possible 'whiffs' of O_2 late in the Archaean; their duration and magnitude are poorly understood. An additional frontier lies in reconstructing the detailed fabric of 'state changes' in atmospheric PO_2, such as occurred at the transitions from the late part of the Archaean to the early Proterozoic and from the late Proterozoic to the early Phanerozoic (*blue boxes*) Reproduced from Lyons et al. (2014). With kind permission of © Nature

(Kasting and Siefert 2002), various models can be proposed for the evolution of abiotic and biotic factors cycles and the microbial lineages (metabolic pathways) capable of driving them (Zerkle et al, 2006; Capone et al. 2006; Navarro-González et al. 2001; Wächtershäuser 2007), but it is clear that oxygenation led to drastic changes in the availability of mineral factors required for metabolic reactions.

The distribution of life on earth is constrained also by the distribution of 20 bio-essential nutrients such as Calcium (Ca), Chloride (Cl^-), Chromium (Cr), Cobalt (Co), Copper (Cu), Magnesium (Mg), Manganese (Mn), Nickel (Ni), Iodine (I), Iron (Fe), **Molybdenum (Mo)**, Phosphorus (P), Potassium (K), Selenium (Se), Sodium (Na), Sulfur (S), Tugsten (W), Vanadium (V), and Zinc (Zn), which are relatively rare, but that are key components of DNA, RNA, enzymes, and other biomolecules. Therefore, the availability of these bio-essential elements must have shaped the evolution of life (Williams and Fraústo da Silva 2003). Particularly, the availability of those elements required as cofactor might have influenced the timing of origin and evolution of metabolic pathways (and hence of phenotypic characteristics). This idea relies on fossil evidences that suggest that the diversification of eukaryotes substantially coincided with the increase in redox potential of the deep oceans about 800 million years ago (Knoll et al. 2006), with a subsequent increase in zinc's, molybdenum's and other elements' availability and the parallel decrease of iron, manganese, and cobalt concentrations. Bioinformatic analyses of protein-metal binding motifs have shown that eukaryotes require more zinc and less iron, manganese and cobalt in respect to prokaryotes (Dupont et al. 2006). Eukaryotes also require molybdenum for nitrate assimilation (Anbar and Knoll 2002) and can use zinc instead of cobalt for assimilation of carbon through the enzyme carbonic anhydrase. It is therefore possible that eukaryotes emerged from ecological niches such as oceanic chemical mass shifted to favor the needs for the elements (Anbar and Knoll 2002; Knoll 2006). A similar pathway could explain the arisal of red eukaryotic phytoplankton after 250 million years ago (Falkowski et al. 2004).

Moreover, for example, the availability of trace metals of importance to the N cycle (Fe, Mo, and Cu) was influenced by increasing concentration of oxygen since they are redox sensitive and display different availability depending on redox state of the environment. Indeed, they exhibit sharp changes in the geochemical behaviour with changes in environmental oxygenation; Fe and Mo, for example, are a complementary pair because they are geochemical opposites. In Earth's crust (Taylor and McLennan 1982), Fe is the fourth most abundant element (~5%), whereas Mo is a trace element (~1 ppm). This situation is reversed in the oceans where Fe concentrations rarely exceed ~2 nM (Bruland et al. 1979; Wu et al. 2001), whereas the concentration of Mo is a uniform ~105 nM (Collier 1985). In oxygenated waters, Fe (III) oxidizes and hydrolyzes forming an insoluble oxyhydroxide whereas Fe (II) is mobile, but is sequestered into sulfides. Molybdenum (VI) forms a soluble oxyanion (MoO_4^{2-}) in oxic water, and is readily transferred to the oceans in which its residence time is ~800,000 years (Morford and Emerson 1999). On the contrary, $MoO_{4-x}S_x^2$ is formed in reduced conditions.

2.2 The History of Molybdenum Availability and Utilisation

Molybdenum (Mo) is an elemental metal of the second transition series, which is adjacent to chromium (Cr) and tungsten (W); Mo occurs in nature with a + 6 oxidation state, its Pauling electro negativity of 2.16 allows it to form covalent bonds. A summary of molybdenum physical and atomic properties is shown in Table 2.3. Molybdenum has a diagonal relationship with vanadium (V) and rhenium (Re) in the periodic table; an example of this is that in their highest oxidation states, these elements form tetrahedral oxyanions VO_4^{3-}, MoO_4^{2-} and ReO_4^{-}, rhenium minerals may be associated with molybdenum disulfide (molybdenite, MoS_2).

Table 2.3 Atomic and physical properties of molybdenum and tungsten

	Molybdenum	Tungsten
Atomic number	42, transition metal	23, transition metal
Atomic mass	95, 94 g amu	50, 9415 amu
Electronic configuration	$[Kr]4d^5 5s^1$	$[Ar]3d^3 4s^2$
Oxidation states	2, 3, 4, 5, 6	5, 3
Density	10,280 kg/m^3	6110 kg/m^3
Covalent radius	145 pm	125 pm
Cristal structure	Body centered cubic	Body centered cubic
Melting point	2896 K	2175 K
Boiling point	4912 K	3682 K
Electronegativity	2, 16	1, 63
Free energy of solvation	−226.8 kcal mol^{-1}	−230.1 kcal mol^{-1}
Covalent solution radii	2, 75 A	2, 83 A

From a geochemically viewpoint, Mo is mainly found in nature in the form of molybdate (MoO_4^{2-}). Mo is so unreactive in oxygenated aqueous solutions that is the most abundant transition metal in the modern oceans (~105 nmol kg^{-1}), with a residence time of ~800 ka (Morford and Emerson 1999), despite being a component whose concentration in the cortex is found only in parts per million (ppm). In contrast, Mo is easily removed from the water column in euxínic conditions, i.e. anoxic environments with high concentrations of H_2S, so that the enrichment of Mo in sediments is considered diagnostic of reducing deposition conditions (Calvert and Pedersen 1993; Anbar 2004; McManus et al. 2006; Tribovillard et al. 2006), in fact where sulfide is present in the water column or in pore fluids, Mo accumulates at a 100–1000 fold higher rate than when sulfide is absent. Thus, the acqueous Mo-cycle and Mo isotope mass balance are obviously affected by these different settings that may occur, i.e. euxinc and oxic conditions.

Molybdenum enters the oceans largely via rivers. In oxygenated waters, Mo speciation is mostly represented by the molybdate anion, MoO_2. However, under euxinic conditions, such as in the modern Black Sea, where the concentration of H_2S is very high (about 100 M) several Mo species are readily removed from solution (Emerson 1991; Erickson and Helz 2000; Morford and Emerson 1999). As a result, ocean sediments accumulating today beneath euxinic waters account for 10 to 50 % of the annual removal flux of Mo from the oceans even though such waters cover an extremely small percentage of the modern seafloor (0.5 %) (Emerson 1991; Morford and Emerson 1999). The remaining marine Mo is primarily removed in association with sedimentary Mn in oxygenated environments (Morford and Emerson 1999). It is not known if sediments accumulating under bottom waters with low O_2 concentration and lacking sulfide ("suboxic" sediments) constitute a net source or sink of Mo from the oceans (Morford and Emerson 1999; McManus 2002), but they probably account for about 10% of Mo cycling today (Morford and Emerson 1999). The ocean Mo budget is therefore highly sensitive to the extent of euxinic bottom waters (Emerson 1991).

2.2.1 Isotope Fractionation in the Modern Ocean

Mo has seven naturally occurring stable isotopes: ^{92}Mo (14.84 %), ^{94}Mo (9.25 %), ^{95}Mo (15.92 %), ^{96}Mo (16.68 %), ^{97}Mo (9.55 %), ^{98}Mo (24.13 %), and ^{100}Mo (9.63 %). Natural mass-dependent Mo-isotope variations span a range of ~1 ‰ amu^{-1} (Atomic Mass Unit) difference, usually reported as $\delta^{97/95}$Mo or $\delta^{98/95}$Mo.

The largest known Mo-isotope effect occurs during adsorption of dissolved Mo to Mn-oxide particles and Fe(oxyhydr)oxides, (Siebert et al. 2003; Barling and Anbar 2004; Wasylenki et al. 2006; Goldberge al. 2009), concentrating isotopically-light Mo in ferromanganese sediments deposited in oxic settings (e.g., Barling et al. 2001). In contrast, in euxinic basins with free sulfide in the water column, Mo is nearly completely removed into the sediments such that no fractionation from the seawater value is expressed (Barling et al. 2001;

Arnold et al. 2004; Nägler et al. 2005). Light isotopes are preferentially removed from solution, so dissolved Mo is ~1 ‰ amu^{-1} heavier than particle-bound Mo.

First inferred from natural observations in oxic seawater and sediments (Barling et al. 2001), this fractionation has been reproduced in the laboratory (Barling and Anbar 2004; Wasylenki et al. 2006). The fractionation factor is relatively unaffected by pH or temperature. Ironically, despite the rich redox chemistry of Mo (Anbar 2004), this fractionation may not involve a change in Mo-oxidation state. It may result from the fractionation between MoO_4^{2-}, which dominates Mo speciation in seawater, and one or more scarce, dissolved Mo species, such as $Mo(OH)_6$ or MoO_3, that adsorbs more readily (Siebert et al. 2003; Tossell 2005). Alternatively, fractionation may result from isotope exchange between MoO_4^{2-} and a surface-bound Mo species (Barling and Anbar 2004).

Smaller but significant Mo-isotope effects are inferred to exist during the reduction of Mo in suboxic and euxinic environments based on observations in reducing sedimentary basins (McManus et al. 2002, 2006; Nägler et al. 2005; Poulson et al. 2006; Siebert et al. 2003). Mo is highly reactive in reduced settings, particularly when dissolved (H_2S) > 11 μmol kg^{-1} and MoO_4^{2-} is transformed to particle-reactive oxythiomolybdates, $MoO_{4-x}S_x^{2-}$ (Erickson and Helz 2000; Helz et al. 1996; Zheng et al. 2000).

Mo isotope variations magnitudes are also seen in continental molybdenites (Barling et al. 2001; Malinovsky et al. 2005; Pietruszka et al. 2006; Wieser and de Laeter 2003). Biological isotope effects appear to be minor, although there are reports of preferential uptake of light isotopes (by ~0.25 ‰ amu^{-1}) by nitrogen-fixing bacteria (Liermann et al. 2005; Nägler et al. 2004).

It follows from the long residence time of Mo in oxic seawater that, in contrast to Fe, Mo-isotope composition of bulk oxic sediments is controlled by fractionation during Mo removal to such sediments, if such fractionation occurs, as well as by the isotopic composition of Mo sources to the oceans. Similar considerations can be applied to suboxic sediments, which are turning out to be important to the Mo ocean budget (McManus et al. 2006). Hence, Mo-isotope oceanographic research has focused at least as much on understanding isotope effects during Mo removal to sediments as on Mo sources to the oceans. Changes in Mo removal fluxes under steady-state conditions are shown in Fig. 2.3.

Early data pointed to a first-order model in which Mo entering the oceans primarily from continental sources with average $\delta^{97/95}$Mo ~ 0 ‰ undergoes fractionation during removal to Mn oxides, resulting in an isotopically heavy reservoir of Mo dissolved in seawater. The removal of Mo to euxinic sediments imparts little isotope effect. From a paleoceanographic perspective, this model implies that $\delta^{97/95}$Mo seawater should vary with the extent of oxic deposition; if there were no Mn-oxide deposition, $\delta^{97/95}$Mo seawater ~0 ‰ (Barling et al. 2001; Siebert et al. 2003). This schematic concept underlies the use of Mo isotopes as paleoredox proxy, but has three major uncertainties. First, continental rocks being weathered may not be isotopically uniform, so changes in weathering regimes could lead to changes in $\delta^{97/95}$Mo seawater independent of ocean redox. Second, the extent of Mo-isotope fractionation during weathering and transport to the oceans

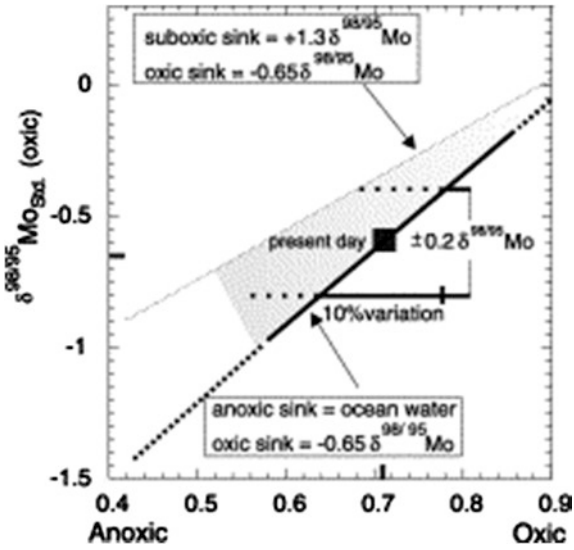

Fig. 2.3 Changing anoxic to oxic removal flux in the oceans under steady-state conditions. The oxic fraction of total Mo removal in the modern oceans must be around 0.7 to maintain steady-state conditions. Variations in the relative amount of oxic and reducing Mo removal flux greater than 10 % should result in a shift in ocean water $\delta^{98/95}$Mo, detectable in the Fe–Mn crust time series. *Shaded area* various mixtures of anoxic (ocean water Mo) to suboxic (lighter isotope compositions) fractionation result in a variety of slopes. A lighter suboxic sink results in less resolution for the detection of variations in the ratio of oxic to anoxic Mo fluxes. Reproduced from Siebert et al. (2003). With kind permission of © Earth and Planetary Science Letters

is not known. Leaching experiments with igneous rocks suggest that weathering effects are minimal (Siebert et al. 2003), but isotopic research in rivers and estuaries is needed. Third, and most important, suboxic sedimentary environments, once thought to be of negligible importance to the Mo elemental and isotope budgets (Barling et al. 2001; Morford and Emerson 1999), may account for as much as 60 % of Mo removal (McManus et al. 2006). If so, an effect on the ocean isotope budget is likely because Mo isotopes are significantly fractionated in pore waters and sediments from such settings (McManus et al. 2002; Poulson et al. 2006). Relative to seawater Mo, the direction of this effect is the same as fractionation by removal to Mn-oxide sediments, so in paleoceanographic reconstructions changes in $\delta^{97/95}$Mo seawater owing to the contraction of oxic deposition could be partially countered by the expansion of suboxic deposition, and *vice versa*.

Researchers are exploring three types of Mo-isotope paleorecords. Temporal variations of $\delta^{97/95}$Mo seawater could be mirrored in Fe-Mn crusts, providing a paleorecord through the Cenozoic (Siebert et al. 2003). An alternative record extending much further in time may exist in black shales (Barling et al. 2001). Finally, Mo isotopes may also be useful in probing local, rather than global, redox conditions.

Siebert et al. (2003) examined Mo isotopes in well-dated Fe-Mn crusts from the Pacific, Atlantic, and Indian Oceans. These sediments span ~60 million years and were sampled at a resolution of ~1–3 million years. The Mo- isotope composition of all these samples is remarkably uniform. Overall, Mo in these sediments varies within a range of $<\pm 0.25$ ‰. The data overlap with measurements of Atlantic and Pacific Fe-Mn nodules (Barling et al. 2001), despite greatly different accumulation rates of these sediment types. The only significant $\delta^{97/95}$Mo variation in these samples is a small offset of ~0.15 ‰ seen between Pacific and Atlantic crusts, just outside the $\pm 2\sigma$ uncertainties of Siebert et al. (2003). The isotopic uniformity of this record is taken to reflect the invariance of $\delta^{97/95}$Mo seawater at the sampled resolution through the Cenozoic and may indicate that ocean oxygenation has varied less than 10 % from modern values over this time (Siebert et al. 2003). Although larger ocean redox perturbations are possible during this time in association with glacial pCO_2 drawdown, their duration does not approach the $\sim 10^6$ year timescale of crust sampling and the mean ocean residence time of Mo.

Since the amount of bio-available Mo strongly changed during evolution to form anoxic to oxic environments, it is interesting to speculate about the appearance of Mo containing enzymes in relation to other metals with similar chemical characteristics.

2.3 Molybdenum Versus Tungsten

It is interesting and intriguing to speculate on the reason why evolution selected Mo and W as the metal component of enzymes cofactor. It is known that in oxic aqueous systems, these metals exist in the form of their tetrahedral MoO_4^{2-} or WO_4^{2-} oxyanions; therefore, they are easily mobilizable into enzyme systems with the important caveat that they must be distinguished during enzyme maturation to ensure proper redox function and therefore catalysis.

But how did these two systems arise? And why have both of them been selected?

It is plausible that, at the beginning, W might have been preferred by the early biological systems since tungsten–sulphur bonds, such as those found in tungsten-containing enzymes, are more stable than their molybdenum counterparts. Moreover, if we agree with the idea of an anaerobic and (highly) reducing environment, low-valent tungsten sulfides are also more soluble in aqueous solutions and, as a result, W would have been more bioavailable than Mo in the primordial Earth's biosphere (Andreesen and Makdessi 2008). Lastly, the reduction potentials of tungsten-containing complexes are lower than those containing molybdenum and therefore more useful to early life forms, which probably had a low intracellular redox poise. To date the most reliable explanation for the existence of both Mo and W containing enzymes is based on the evolution of life on Earth along with the subsequent biogeochemistry impact of (micro)organisms on the biosphere (Hille 2002; Dietrich et al. 2006). In fact,

as organisms performing oxygenic photosynthesis rendered the atmosphere progressively aerobic, the O_2 sensitivity of tungsten compounds became a liability and the water-solubility of high-valent molybdenum oxides an asset; the elevated intracellular redox poise of aerobic organisms would have also made molybdenum more suitable. Following the evolution of biological water oxidation, oxygenic photosynthetic organisms caused a dramatic increase in ambient redox potential that paralleled the increase in atmospheric oxygen, and this resulted in the appearance and bioavailability of Mo in the form of MoO_4^{2-}. As a consequence, the bioavailability of the two elements was reversed, with Mo being present at a concentration in seawater at 100 times that of W (Anbar 2008). Quite interestingly, the processes that rendered Mo bioavailable had an opposite effect on iron bioavailability; indeed the increasing oxygen concentrations allowed to oxidize iron(II) to iron(III) oxyhydroxides, which are (essentially) insoluble and hence less bioavailable (Anbar and Knoll 2002; Fike et al. 2006). Since tungsten and molybdenum have similar chemistry, it is possible that, initially, as the transition to an oxygen-rich environment occurred, the latter progressively substituted for the former in enzyme active sites. Indeed, molybdenum and tungsten containing enzymes often catalyze the same kind of reactions, such as the oxidation of aldehydes to carboxylic acids. This idea is supported by the biological distribution of these elements today: molybdenum containing enzymes are found in all aerobic organisms, whereas tungsten occurs only in obligate, typically thermophilic, anaerobes. Accordingly, some mesophilic anaerobes are more "plastic" than their aerobic counterparts, in the sense that they can utilize either molybdenum or tungsten depending on their availability and growth conditions. However, both Mo and W are able to assemble into mononuclear molybdoenzymes in an almost identical manner causing a biological problem since redox reactions catalyzed by W typically occur at much lower potentials than those catalyzed by Mo, and the active sites of the enzymes have evolved to modulate the redox properties of their cognate metal ion. Therefore, the replacement of Mo by W in a *bona fide* molybdoenzyme would elicit dramatic changes in catalytic efficacy. This has been demonstrated using W substituted dimethyl sulfoxide reductase from *Rhodobacter capsulatus* in which the (V/VI) and (IV/V) redox potentials were decreased by −220 and −334 mV, respectively (Hagedoorn et al. 2003). The W substituted enzyme is able to catalyze dimethyl sulfoxide reduction (the forward reaction), but cannot catalyze dimethylsulfide oxidation (the backward reaction) (Hagedoorn et al. 2003). In the *Escherichia coli* trimethylamine-N-oxide reductase, substitution of Mo with W also results in retention of catalysis in the forward direction while considerably perturbing its substrate specificity (Buc et al. 1999). In the case of sulfite oxidase, W substitution results in completely inactive enzyme (Johnson and Rajagopalan 1977). In the aldehyde: ferredoxin oxidoreductase of the Archaeon *Pyrococcus furiosus*, substitution of Mo for W results in enzyme unable to catalyze aldehyde oxidation (Sevcenco et al. 2010). Toxicity of the antagonist oxyanion *via* incorrect metal insertion occurs in both W toward molybdoenzymes and Mo toward tungstoenzymes.

Bibliography

Anbar, A. D. (2004). Molybdenum stable isotopes: Observations, interpretations and directions. *Reviews in Mineralogy and Geochemistry, 55*(1), 429–454.

Anbar, A. D. (2008). Elements and evolution. *Science, 322*, 1481–1483.

Anbar, A. D., & Knoll, A. H. (2002). Proterozoic ocean chemistry and evolution: A bioinorganic bridge? *Science, 297*(5584), 1137–1142.

Andreesen, J. R., & Makdessi, K. (2008). Tungsten, the surprisingly positively acting heavy metal element for prokaryotes. *Annals of the New York Academy of Sciences, 1125*(1), 215–229.

Archer, C., & Vance, D. (2008). The isotopic signature of the global riverine molybdenum flux and anoxia in the ancient oceans. *Nature Geoscience, 1*(9), 597–600.

Arnold, G. L., Anbar, A. D., Barling, J., & Lyons, T. W. (2004). Molybdenum isotope evidence for widespread anoxia in mid-Proterozoic oceans. *Science, 304*(5667), 87–90.

Barling, J., & Anbar, A. D. (2004). Molybdenum isotope fractionation during adsorption by manganese oxides. *Earth and Planetary Science Letters, 217*(3), 315–329.

Barling, J., Arnold, G. L., & Anbar, A. D. (2001). Natural mass-dependent variations in the isotopic composition of molybdenum. *Earth and Planetary Science Letters, 193*(3), 447–457.

Bekker, A., Holland, H. D., Wang, P. L., Rumble, D. I. I. I., Stein, H. J., Hannah, J. L., & Beukes, N. J. (2004). Dating the rise of atmospheric oxygen. *Nature, 427*(6970), 117–120.

Böning, C. W., Dispert, A., Visbeck, M., Rintoul, S. R., & Schwarzkopf, F. U. (2008). The response of the Antarctic circumpolar current to recent climate change. *Nature Geoscience, 1*(12), 864–869.

Bruland, K. W., Franks, R. P., Knauer, G. A., & Martin, J. H. (1979). Sampling and analytical methods for the determination of copper, cadmium, zinc, and nickel at the nanogram per liter level in sea water. *Analytica Chimica Acta, 105*, 233–245.

Buc, J., Santini, C. L., Giordani, R., Czjzek, M., Wu, L. F., & Giordano, G. (1999). Enzymatic and physiological properties of the tungsten-substituted molybdenum TMAO reductase from Escherichia coli. *Molecular Microbiology, 32*(1), 159–168.

Calvert, S. E., & Pedersen, T. F. (1993). Geochemistry of recent oxic and anoxic marine sediments: Implications for the geological record. *Marine Geology, 113*(1), 67–88.

Canfield, D. E. (2005). The early history of atmospheric oxygen: Homage to Robert M. Garrels. *Annual Review of Earth and Planetary Sciences, 33*, 1–36.

Capone, D. G., Popa, R., Flood, B., & Nealson, K. H. (2006). Geochemistry. Follow the nitrogen. *Science, 312*(5774), 708–709.

Collier, R. W. (1985). Molybdenum in the northeast Pacific Ocean. *Limnology and Oceanography, 30*(6), 1351–1354.

Crusius, J., Calvert, S. E., Pedersen, T. F., & Sage, D. (1996). Rhenium and molybdenum enrichments in sediments as indicators of oxic, suboxic and anoxic conditions of deposition. *Earth and Planetary Science Letters, 145*, 65–79.

Dietrich, H. M., Törnroos, K. W., & Anwander, R. (2006). "Ionic carbenes": Synthesis, structural characterization, and reactivity of rare-earth metal methylidene complexes. *Journal of the American Chemical Society, 128*(29), 9298–9299.

Dupont, C. L., Yang, S., Palenik, B., & Bourne, P. E. (2006). Modern proteomes contain putative imprints of ancient shifts in trace metal geochemistry. *Proceedings of the National Academy of Sciences, 103*(47), 17822–17827.

Emerson, S. R., & Huested, S. S. (1991). Ocean anoxia and the concentrations of molybdenum and vanadium in seawater. *Marine Chemistry, 34*(3), 177–196.

Erickson, B. E., & Helz, G. R. (2000). Molybdenum (VI) speciation in sulfidic waters: stability and lability of thiomolybdates. *Geochimica et Cosmochimica Acta, 64*(7), 1149–1158.

Falkowski, P. G., Fenchel, T., & Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science, 320*(5879), 1034–1039.

Falkowski, P. G., Katz, M. E., Knoll, A. H., Quigg, A., Raven, J. A., Schofield, O., & Taylor, F. J. R. (2004). The evolution of modern eukaryotic phytoplankton. *Science, 305*(5682), 354–360.

Fike, D. A., Grotzinger, J. P., Pratt, L. M., & Summons, R. E. (2006). Oxidation of the Ediacaran ocean. *Nature, 444*(7120), 744–747.

Goldberg, T., Archer, C., Vance, D., & Poulton, S. W. (2009). Mo isotope fractionation during adsorption to Fe (oxyhydr) oxides. *Geochimica et Cosmochimica Acta, 73*(21), 6502–6516.

Hagedoorn, P. L., Hagen, W. R., Stewart, L. J., Docrat, A., Bailey, S., & Garner, C. D. (2003). Redox characteristics of the tungsten DMSO reductase of *Rhodobacter capsulatus*. *FEBS Letters, 555*(3), 606–610.

Helz, G. R., Miller, C. V., Charnock, J. M., Mosselmans, J. F. W., Pattrick, R. A. D., Garner, C. D., & Vaughan, D. J. (1996). Mechanism of molybdenum removal from the sea and its concentration in black shales: EXAFS evidence. *Geochimica et Cosmochimica Acta, 60*(19), 3631–3642.

Hille, R. (2002). Molybdenum and tungsten in biology. *Trends in Biochemical Sciences, 27*(7), 360–367.

Holland, H. D. (2006). The oxygenation of the atmosphere and oceans. *Philosophical Transactions of the Royal Society B: Biological Sciences, 361*(1470), 903–915.

Isley, A. E., & Abbott, D. H. (1999). Plume-related mafic volcanism and the deposition of banded iron formation. *Journal of Geophysical Research: Solid Earth (1978–2012), 104*(B7), 15461–15477.

Johnson, J. L., & Rajagopalan, K. V. (1977). Tryptic cleavage of rat liver sulfite oxidase. Isolation and characterization of molybdenum and heme domains. *Journal of Biological Chemistry, 252*(6), 2017–2025.

Kasting, J. F. (1993). Earth's early atmosphere. *Science, 259*(5097), 920–926.

Kasting, J. F., & Siefert, J. L. (2002). Life and the evolution of Earth's atmosphere. *Science, 296*(5570), 1066–1068.

Knoll, A. H., Javaux, E. J., Hewitt, D., & Cohen, P. (2006). Eukaryotic organisms in Proterozoic oceans. *Philosophical Transactions of the Royal Society B: Biological Sciences, 361*(1470), 1023–1038.

Lazcano, A., & Miller, S. L. (1996). The origin and early evolution of life: Prebiotic chemistry, the pre-RNA world, and time. *Cell, 85*(6), 793–798.

Liermann, L. J., Guynn, R. L., Anbar, A., & Brantley, S. L. (2005). Production of a molybdophore during metal-targeted dissolution of silicates by soil bacteria. *Chemical Geology, 220*(3), 285–302.

Lyons, T. W., Reinhard, C. T., & Planavsky, N. J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature, 506*, 307–315.

Malinovsky, D., Rodushkin, I., Baxter, D. C., Ingri, J., & Öhlander, B. (2005). Molybdenum isotope ratio measurements on geological samples by MC-ICPMS. *International Journal of Mass Spectrometry, 245*(1), 94–107.

McManus, J., Berelson, W. M., Severmann, S., Poulson, R. L., Hammond, D. E., Klinkhammer, G. P., & Holm, C. (2006). Molybdenum and uranium geochemistry in continental margin sediments: Paleoproxy potential. *Geochimica et Cosmochimica Acta, 70*(18), 4643–4662.

McManus, J., Nägler, T. F., Siebert, C., Wheat, C. G., & Hammond, D. E. (2002). Oceanic molybdenum isotope fractionation: Diagenesis and hydrothermal ridge-flank alteration. *Geochemistry, Geophysics, Geosystems, 3*(12), 1–9.

Morford, J. L., & Emerson, S. (1999). The geochemistry of redox sensitive trace metals in sediments. *Geochimica et Cosmochimica Acta, 63*(11), 1735–1750.

Nägler, T. F., Mills, M. M., & Siebert, C. (2004). Biological fractionation of Mo isotopes during N2 fixation by Trichodesmium sp. IMS 101. *Geochimica et Cosmochimica Acta, 68*, A364.

Nägler, T. F., Siebert, C., Lüschen, H., & Böttcher, M. E. (2005). Sedimentary Mo isotope record across the Holocene fresh–brackish water transition of the Black Sea. *Chemical Geology, 219*(1), 283–295.

Navarro-González, R., McKay, C. P., & Mvondo, D. N. (2001). A possible nitrogen crisis for Archaean life due to reduced nitrogen fixation by lightning. *Nature, 412*(6842), 61–64.

Neubert, N., Nägler, T. F., & Böttcher, M. E. (2008). Sulfidity controls molybdenum isotope fractionation into euxinic sediments: Evidence from the modern Black Sea. *Geology, 36*(10), 775–778.

Pearce, C. R., Cohen, A. S., Coe, A. L., & Burton, K. W. (2008). Molybdenum isotope evidence for global ocean anoxia coupled with perturbations to the carbon cycle during the early Jurassic. *Geology, 36*(3), 231–234.

Pietruszka, A. J., Walker, R. J., & Candela, P. A. (2006). Determination of mass-dependent molybdenum isotopic variations by MC-ICP-MS: An evaluation of matrix effects. *Chemical Geology, 225*(1), 121–136.

Poulson, R. L., Siebert, C., McManus, J., & Berelson, W. M. (2006). Authigenic molybdenum isotope signatures in marine sediments. *Geology, 34*(8), 617–620.

Poulton, S. W., Fralick, P. W., & Canfield, D. E. (2004). The transition to a sulphidic ocean ∼1.84 billion years ago. *Nature, 431*(7005), 173–177.

Prantzos, N., Vangioni-Flam, E., & Chauveau, S. (1994). Evolution of carbon and oxygen in the galaxy: The effect of metallicity dependent yields. *Astronomy and Astrophysics, 285*, 132–139.

Rubey, W. W. (1955). Development of the hydrosphere and atmosphere, with special reference to probable composition of the early atmosphere. *Geological Society of America Special Papers, 62*, 631–650.

Sageman, B. B., & Lyons, T. W. (2003). Geochemistry of fine-grained sediments and sedimentary rocks. *Treatise on Geochemistry, 7*, 115–158.

Sevcenco, A. M., Bevers, L. E., Pinkse, M. W., Krijger, G. C., Wolterbeek, H. T., Verhaert, P. D., et al. (2010). Molybdenum incorporation in tungsten aldehyde oxidoreductase enzymes from *Pyrococcus furiosus*. *Journal of Bacteriology, 192*(16), 4143–4152.

Siebert, C., Nägler, T. F., von Blanckenburg, F., & Kramers, J. D. (2003). Molybdenum isotope records as a potential new proxy for paleoceanography. *Earth and Planetary Science Letters, 211*(1), 159–171.

Taylor, S. R., & McLennan, S. M. (1982). Geochemical constraints on the growth of the continental crust. *The Journal of Geology, 90*, 347–361.

Tossell, J. A. (2005). Calculating the partitioning of the isotopes of Mo between oxidic and sulfidic species in aqueous solution. *Geochimica et Cosmochimica Acta, 69*(12), 2981–2993.

Tribovillard, N., Algeo, T. J., Lyons, T., & Riboulleau, A. (2006). Trace metals as paleoredox and paleoproductivity proxies: An update. *Chemical Geology, 232*(1), 12–32.

Wächtershäuser, G. (2007). On the chemistry and evolution of the pioneer organism. *Chemistry and Biodiversity, 4*(4), 584–602.

Wasylenki, L. E., Anbar, A. D., & Gordon, G. W. (2006). Temperature dependence of Mo isotope fractionation during adsorption to δ-MnO_2: Implications for the paleoredox proxy. *Geochimica et Cosmochimica Acta, 70*(18), A691.

Wieser, M. E., & de Laeter, J. R. (2003). A preliminary study of isotope fractionation in molybdenites. *International Journal of Mass Spectrometry, 225*(2), 177–183.

Williams, R. P., & Fraústo da Silva, J. J. R. (2003). Evolution was chemically constrained. *Journal of Theoretical Biology, 220*(3), 323–343.

Wu, J., Boyle, E., Sunda, W., & Wen, L. S. (2001). Soluble and colloidal iron in the oligotrophic North Atlantic and North Pacific. *Science, 293*(5531), 847–849.

Zerkle, A. L., House, C. H., Cox, R. P., & Canfield, D. E. (2006). Metal limitation of cyanobacterial N_2 fixation and implications for the precambrian nitrogen cycle. *Geobiology, 4*(4), 285–297.

Zheng, Y., Anderson, R. F., van Geen, A., & Kuwabara, J. (2000). Authigenic molybdenum formation in marine sediments: A link to pore water sulfide in the Santa Barbara Basin. *Geochimica et Cosmochimica Acta, 64*(24), 4165–4178.

Chapter 3
Molybdenum and Biological Systems (Molybdenum Cofactors Containing Enzymes and Pathways)

Abstract Molybdenum and tungsten share several atomic properties. Hence, in order to distinguish them, microorganisms have evolved discriminating systems either at the level of metal insertion into the cofactor or during the possible interplay of molybdo-enzyme specific chaperones when incorporating the cofactor. Several different Mo transport systems have been identified in Prokaryotes and Eucaryotes with their relatively different regulation and storage systems.

Molybdenum and tungsten have very similar atomic radii (1.75 and 1.78 A, respectively), similar electronegativity, the same free energy of solvation (−226.8 and −230.1 kcal mol^{-1}, respectively) and the same covalent solution radii (2.75 and 2.83 A, respectively), other atomic properties are showed in Table 2.3. Hence, in order to distinguish them, microorganisms have evolved discriminating systems at the level of metal insertion into the cofactor (Nichols and Rajagopalan 2002; Nichols et al. 2007; Bevers et al. 2008) and the possible interplay of molybdoenzyme specific chaperones with cofactor incorporation (Blasco et al. 1998; Chan et al. 2006; Feng et al. 2009).

3.1 Molybdenum Uptake in Bacteria

In Bacteria, molybdenum is transported into the cytoplasm in the form of the oxyanion molybdate (MoO_4^{2-}), which is taken up through the membrane by high-affinity ABC-type transporters (Aguilar-Barajas et al. 2011). These transporters allow bacterial cells to scavenge the less abundant molybdate in the presence of sulfate, whose concentration in seawater is about 105 times higher than that of molybdate (Schwarz et al. 2007). Remarkably and not surprisingly, due to the similarity they share, molybdate and tungstate oxyanions may be transported by the same carrier (Grunden and Shanmugam 1997; Bevers et al. 2006).

L. Presta et al., *Molybdenum Cofactors and Their role in the Evolution of Metabolic Pathways*, SpringerBriefs in Biometals, DOI 10.1007/978-94-017-9972-0_3

At least three different Mo transport systems have been identified in Bacteria, namely the Mod, Wtp (W transport protein) and Tup (Tungsten uptake) systems. Based on global sequence comparative analysis, these three systems form distinct entities (Bevers et al. 2006) and, consequently, it is impossible to retrace any putative evolutionary relationships between them.

One of the consequences of the lack of shared sequence similarity is their selectivity in their affinity for molybdate versus tungstate. Indeed, while Tup and Wtp systems are strongly selective in favor of tungstate, Mod system does not discriminate between the two oxoanions. However, despite their differences at the level of the primary sequence, the various molybdated/tungstate binding proteins seem to share the same overall tridimensional structure (Hagen 2011). This is the case, for example, of the overall structure of *Escherichia coli* ModA and *Archeoglobulus fulgidus* WtpA that are quite similar in terms of alpha-helical and beta-sheet elements and in the position of the oxoanion binding site (approximately in the protein's center of gravity (Hollenstein et al. 2009).

3.2 The Mod Transporter System

The Mod complex (Fig. 3.1), which binds ATP to use its Gibbs free energy of hydrolysis for active transport, is made of three different subunits (ModA, ModB and ModC) whose encoding genes are organised in a three-cistronic operon (formerly known as the *chl*D locus) (see below). The ModA protein binds Mo in the periplasm, ModB is a transmembrane component of the permease, and ModC provides the energizer function on the cytoplasmic side of the membrane. Their structure and functioning has been mainly studied in *E. coli* (Hu et al. 1997), *Azotobacter vinelandii* (Balan et al. 2008) and *Xanthomonas axonopodis citri* (Balan et al. 2008). In *E. coli* the successful production of molybdoenzymes relies upon the efficient uptake of molybdate via the molybdate-septic transporter encoded by the *modABC* operon. The analysis of the DNA sequence of the *mod* locus genes revealed that this cluster is constituted by two divergent operons (Fig. 3.2), the first one embedding the three above mentioned genes (*modABC*); the second operon is a bicistronic one constituted by *modE* and *modF*. The structural genes for the transport process are encoded by the *modABC* genes that are transcribed together into a single three-cistronic mRNA, whereas a regulatory protein (ModE) is encoded by the diverging *modEF* operon located upstream of the *modABC* operon. This molybdate transport machinery closely resembles the established ATP-binding cassette (ABC) transporter motif (Higgins 1992; Maupin-Furlow et al. 1995). As shown in Fig. 3.1, its quaternary structure is presumably pentameric, AB_2C_2.

The *E. coli* molybdate-responsive ModE protein negatively controls the transcription of the *modABC* operon (Mouncey et al. 1995; Walkenhorst et al. 1995; Grunden and Shanmugam 1997).

The *modA* gene codes for a 257 amino acid precursor protein from which a 24 amino acid signal peptide is cleaved off to produce a 233 amino acid mature protein.

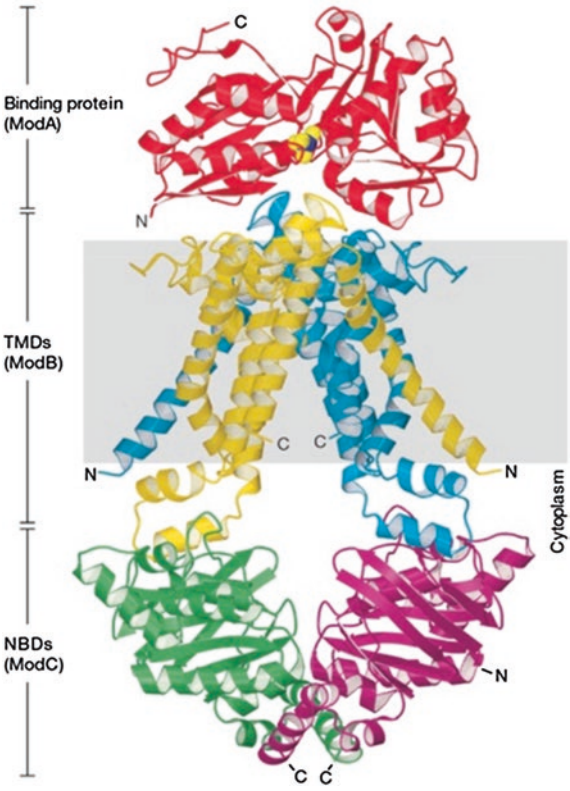

Fig. 3.1 Front view of the $ModB_2C_2$ A complex in ribbon representation, with the ModB subunits coloured *yellow* and *blue*, the ModC subunits coloured *green* and *magenta*, the binding protein ModA coloured *red*, and with bound tungstate in van der Waals representation (*yellow* and *blue spheres*). The *grey box* depicts the probable location of the lipid bilayer on the basis of the hydrophobicity of the protein surface. *N* amino terminus; *C* carboxy terminus. Note that there is a vertical two-fold molecular and non-crystallographic symmetry axis for $ModB_2C_2$. Reproduced from Hollenstein et al. (2007). With kind permission of © Nature

The periplasmic binding protein ModA binds molybdate as well as its analog tungstate at equimolar quantities with a KD of 20 ± 8 nM (Rech et al. 1995, 1996; Imperial et al. 1998). Concerning the structure of the ModA protein, it consists of two well-separated globular domains connected by a hinge region (Fig. 3.1). Upon ligand binding, the two domains interact creating a deep cleft in which the ligand interfaces the two domains. In the absence of bound ligand, the cleft is open and is accessible to bulk solvent (Self et al. 2001). Moreover, the crystal structures that have been determined for ModA complexed with its substrate from a wide range of microorganisms (Chao et al. 1997; Lawson et al. 1998; Balan et al. 2008), allowed the comprehension of the details of its binding mode. As shown in Fig. 3.3 seven atoms of the ModA protein are in H-bond donating distance from the molybdate oxygens: three alcoholic oxygen groups (two Ser and one Tyr) and four nitrogen backbones, including two from the Ser's. There is no indication for significant distortion of the molybdate (or tungstate) tetrahedron (Hagen 2011). From a comparative analysis of 67 ModA-related sequences, the overall organization of the Mo binding site as well as the number of hydrogen bonds appear to be conserved (Hagen 2011).

The ModB protein is a 229-aminoacids long (24,000 Da) and contains six hydrophobic regions, suggesting that the primary location of this protein is the cellular membrane, with a potential sixth region in the C-terminal part of the

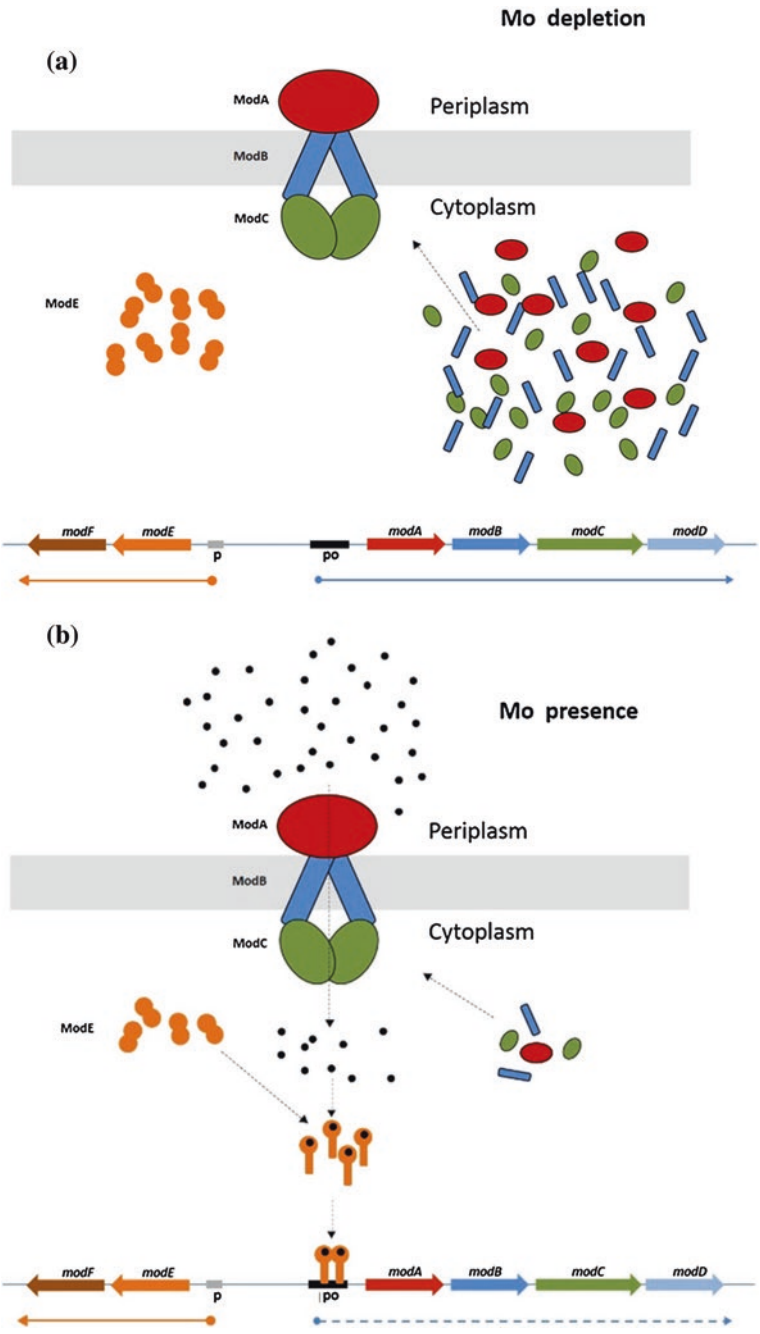

Fig. 3.2 *modABC*-regulation. In the absence of molybdenum the *modABC* genes are transcribed and the corresponding proteins are synthesized (**a**). In the presence of molybdenum (**b**) the ModE-Mo-complex is formed, binds to DNA to inhibit the transcription of the downstream genes *modABC*

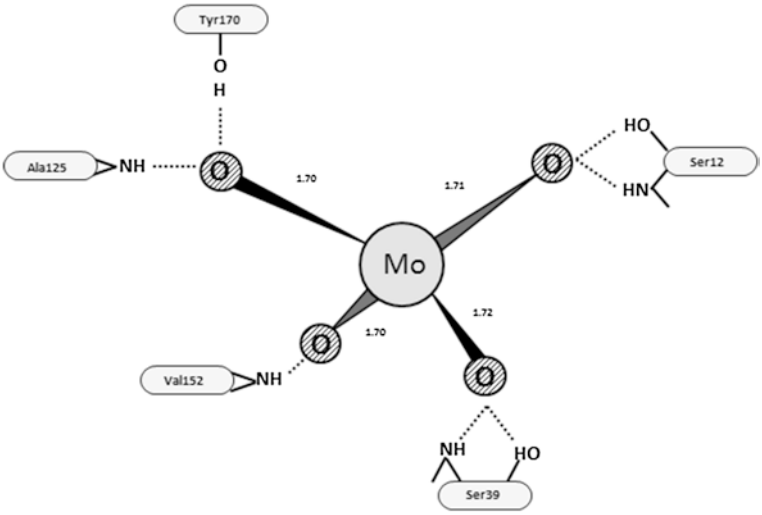

Fig. 3.3 A schematic view of the Mo coordination in the ModA *E. coli* protein (modified from Hagen 2011)

protein. The presence of six distinct hydrophobic domains is reminiscent of the domain structure found in membrane proteins from other members of the ABC-transporter family. With its two TransMembrane Domains (TMDs) that form a translocation pathway for the substrate, ModB provides a channel through which molybdate traverses across the membrane (Self et al. 2001).

The third protein coded for by *modABC* operon is ModC (352-amino acid long in *E. coli*), which possesses the Walker A and B motifs as well as the ABC signature sequence associated with the ABC-ATPases (Holland and Blight 1999). It has two cytoplasmic Nucleotide-Binding Domains (NBDs) that bind and hydrolyze ATP and power the transport reaction (Gerber et al. 2008).

Soon downstream from the *modC* gene, a 696-bp-long ORF has been detected. This ORF encodes a protein, ModD, of 231 amino acids with a molecular mass of 26,336 Da; its amino acid sequence is unique and no protein showing a significant degree of sequence similarity has been reported to date and its function is still unknown (Hagen 2011).

Concerning the working mechanisms of the Mo transport, useful hints have been obtained through X-ray crystallography studies mainly conduced on the Archaea *Methanosarcina acetivorans* and *A. fulgidus* (Hollenstein et al. 2007; Gerber et al. 2008). Overall, from these conformational studies, a two-state minimal model for the working mechanism of these ABC transporters has been proposed; binding of ATP might afford an outward-facing conformation which, upon hydrolysis of ATP, returns to an inward-facing conformation (Hollenstein et al. 2007; Gerber et al. 2008; Hagen 2011). In the absence of this nucleotide, the NBDs (the Nucleotide Binding Domains of ABC transporters, i.e. the ModC subunits of the ModABC system) adopt an open conformation. Upon binding ATP, the NBDs adopt a 'closed' conformation, with bound nucleotide sandwiched

inside the ModC subunits (Dawson et al. 2007). This simple two-state scheme, exemplified by the structure of ModBC-A, can, in principle, explain the binding protein-dependent uptake of essential nutrients by ABC importers. In the ATP-bound, outward-facing state, ABC importers can accept substrate from their cognate binding proteins, whereas, in the nucleotide-free, inward-facing state, ABC importers might release substrate that has previously entered the translocation pathway (Dawson et al. 2007). Recently, it has been argued that the crystallographic studies, have not yet led to consistent insights in their mechanism of action (Hagen 2011). Accordingly, several important mechanistic issues still remain to be explored including the stroke/leverage event in these transport models, the order of events in the release of phosphate and molybdate/tungstate and/or the presence of specific binding interactions (Hagen 2011). Thus, it has become increasingly clear that biochemical, biophysical and modelling studies (including the integration of crystallography with multiple forms of time-resolved spectroscopy) will be required to move from the snapshot pictures provided by the structures to a more complete understanding of the dynamics of these proteins (Gerber et al. 2008).

3.3 Regulation of Mod Transporter System

The details of the molecular mechanisms controlling the Mo uptake have been studied in *E. coli.* In this organism, the molybdate transport genes *modABC* are expressed at extremely low levels in the presence of molybdate, whereas their expression is strongly increased under molybdate starvation (Rech et al. 1995; Rosentel et al. 1995); thus, the intracellular levels of molybdate transcriptionally regulate the *modABC* operon in a process mediated by the ModE protein (Fig. 3.2).

The *E. coli* ModE protein (262 amino acid) is a member of the LysR family of transcriptional regulators. The analyses of the ModE crystal structure revealed that it contains two major domains: (i) the N-terminal one of up to amino acid 121 (domain I) forms a winged helix-turn-helix motif and interacts with DNA, and (ii) the C-terminal one (domain II), representing the putative molybdate-binding component and can be further separated into two subgroups each of which forms a β-barrel (Self et al. 2001). ModE works as a homodimer; indeed, upon binding molybdate (with high affinity, Kd = 0.8 lM), the conformation of the protein changes, promoting the dimerization of ModE-molybdate complex(ModE-Mo). ModE may also bind tungstate (Anderson et al. 1997). The protein dimer is the active form of the complex and is able to bind an operator DNA, upstream of the *modABC* operon, which contains an 8-base inverted repeat 5′-TAACGTTA-3′ flanked by two CAT boxes (Grunden et al. 1996; Anderson et al. 1997) (Fig. 3.2) acting as a repressor of the molybdate transporter; moreover, the same complex acts as an enhancer of several genes involved in the synthetic pathways of *molybdenum*-dependent enzymes, in the presence of molybdate.

ModE is not ubiquitously distributed in bacteria. Indeed, a BLAST (Altschul et al. 1990) probing of NCBI microbial genomes database, using the *E. coli* ModE amicoacid sequence as query, revealed the presence of this protein only in a subset of genomes embedding the ModA sequence. Accordingly, it can be surmised that in some microorganisms other than *E. coli*, Mo uptake might be regulated through a different mechanism.

An intriguing alternative to the *E. coli* system is represented by the Mo transport system of the archaeon *M. acetivorans*. In this microorganism, the binding protein–dependent ABC transporter specific for molybdate/tungstate consists of two transmembrane domains (TMDs, ModB subunits) that form a translocation pathway for the substrate, and two cytoplasmic nucleotide-binding domains (NBDs, ModC subunits) that bind and hydrolyze adenosine triphosphate (ATP) and power the transport reaction (Gerber et al. 2008). What distinguishes *M. acetivorans* ModBC from the other molybdate/tungstate transporter ModBC from *A. fulgidus* (Hollenstein et al. 2007) is the presence of a regulatory domain appended to the C terminus of the NBDs. Moreover, unlike *E. coli* Mo transport system (see above), *M. acetivorans* lacks any ModE homologous sequence. Remarkably, in the *M. acetivorans* Mo transporter, binding of molybdate keeps the catalytic motifs of the NBDs separated, thus disrupting the ATP hydrolysis cycle. High concentrations of free cytoplasmic molybdate will thus lead to an inhibition of the transport activity (and hence to a decrease in molybdate uptake) through both processes. This suggests that trans-inhibition can substitute for the genetic regulation of the molybdate transporter. The genome of other organisms such as *E. coli*, *Photorhabdus luminescens*, and *Klebsiella pneumoniae* do contain genes coding for ModE and at the same time feature regulatory domains attached to their ModC subunits, which suggests that molybdate import in these organisms may be regulated by at least two different molecular mechanisms, i.e. both by trans-inhibition and by genetic control of the expression levels (Gerber et al. 2008).

Notably, the regulatory domains of *M. acetivorans* ModC (residues 231–348) and *E. coli* ModE (residues 124–262) are structurally similar (Gerber et al. 2008).

The architectural similarity of the regulatory domains of ModE and *M. acetivorans* ModC creates an unexpected structural link between genetic regulation and trans-inhibition in ABC transporters. In both processes, tight dimers of regulatory domains with two oxyanions (molybdate or tungstate) sandwiched at the shared interface affect the function of the fused partner domains. ModE exploits this to promote the binding of its N-terminal domains to DNA (23), whereas in the ABC transporter *M. acetivorans* ModBC, binding of molybdate keeps the catalytic motifs of the NBDs separated, thus disrupting the ATP hydrolysis cycle. High concentrations of free cytoplasmic molybdate will thus lead to an inhibition of the transport activity (and hence to a decrease in molybdate uptake) through both processes (Gerber et al. 2008).

3.4 Molybdenum Storage

One of the most known and studied enzymes using Mo as cofactor is nitrogenase (see Chap. 5), which is responsible for the biological nitrogen fixation, i.e. the transformation of atmospheric molecular nitrogen into ammonium. As it will be shown in the following chapters, nitrogen fixation has a narrow phylogenetic distribution in that it is restricted only to some prokaryotes (Bacteria and Archaea). One of the most studied nitrogen-fixing microorganism (diazotrophs) is *Azotobacter vinelandii*. It has been reported that this bacterium can accumulate 25 times more molybdenum than that required for maximum nitrogenase activity (Shah 1984). This ability seems to be due to the unique molybdenum storage protein (MoSto) expressed by this organism. The MoSto protein is an α3β3 hexamer of the *mosA* (*Avin43200*) and *mosB* (*Avin43210*) gene products that can store up to 100 molybdenum atoms per hexamer (Fenske et al. 2005). Metal storage within MoSto occurs in the form of a variety of compact polynuclear oxoanions and its incorporation into the protein is a nucleotide-dependent process (Allen et al. 1999) while the release of molybdate from MoSto is ATP-independent. Further, the release appears to be pH-dependent and to occur stepwise, suggesting the involvement of several amino acid groups in the release mechanism. Although MoSto expression is not controlled by *nif* (nitrogen fixation) regulatory factors (see Chap. 5), the level of molybdenum incorporation into MoSto increases in nitrogen fixing cells, thus suggesting that MoSto could act as a molybdenum reservoir destined to FeMo-co synthesis (Hernandez et al. 2009).

3.5 Molybdenum Uptake in Eukaryotes

In contrast to prokaryotes, molybdate transporters in eukaryotes are almost unknown and, to date, genes similar to the bacterial molybdate channel *modB* have not been found in sequenced genomes. Although eukaryotic Mo uptake mechanisms are very poorly understood it can be said that anions sharing some physicochemical characteristics might be transported by carriers from related families; accordingly, for anions such as molybdate, sulfate, and selenite, which are similarly shaped, there are molybdenum transporters found as sulfur/molybdenum cotransporters and phosphate/molybdenum cotransporters (Tejada-Jiménez et al. 2007).

Physiological data from the green alga *Chlamydomonas reinhardtii* suggest the presence of at least two molybdate transport systems that are related to the unlinked genetic *loci nit5* and *nit6* (Llamas et al. 2000). Mutants defective at one of these *loci* show a reduced molybdate transport activity; double mutants at both *loci* lack Mo-co and, thus, activity of the molybdo-enzyme nitrate reductase (González-Ballester et al. 2004). This finding led to the discovery of MoT1 (molybdate transporter, type 1) gene that encodes a protein with homologs both

in eukaryotes and prokaryotes; these proteins share highly conserved motifs that define a previously uncharacterized family of transporters probably involved in molybdate uptake. Several experiments have been carried out to clarify its action mechanism. In 2007 Tejada-Jiménez and coworkers have carried out an expression silencing of the *C. reinhardtii* gene MoT1 showing that *Chlamydomonas* strains with reduced expression of MoT1 exhibited a reduced molybdate transport and nitrate reductase activities, pointing to a molybdate transport function of MoT1. A K_m of 20 nM was obtained for this protein that is comparable with those reported for the molybdate ABC transport system from *E. coli* (50 nM). This K_m value indicates a very high affinity for molybdate, ensuring the provision of an essential micronutrient even if the element is not abundant in the surrounding environment.

It has also been demonstrated that nitrate is a strong activation signal of MoT1 activity, whereas the presence of molybdate does not seem to affect it appreciably (differently from bacteria whose transcriptional level of molybdate transporters depends on intracellular levels of Mo), indicating that regulation of the molybdate transport process mediated by MoT1 might occur mostly at the transcriptional level. The role of nitrate in MoT1 activation points to a direct connection between molybdate transport and nitrate assimilation and corresponding to the Mo requirement for an efficient nitrate reduction. It is remarkable that in addition to this putative molybdate transporter, two other genes for Mo-co biosynthesis, Cnx2 and Cnx6 from *Arabidopsis thaliana*, also respond rapidly to nitrate (Wang et al. 2003).

Bibliography

Aguilar-Barajas, E., Díaz-Pérez, C., Ramírez-Díaz, M. I., Riveros-Rosas, H., & Cervantes, C. (2011). Bacterial transport of sulfate, molybdate, and related oxyanions. *BioMetals, 24*(4), 687–707.

Allen, R. M., Roll, J. T., Rangaraj, P., Shah, V. K., Roberts, G. P., & Ludden, P. W. (1999). Incorporation of molybdenum into the iron-molybdenum cofactor of nitrogenase. *Journal of Biological Chemistry, 274*(22), 15869–15874.

Altschul, S. F., Myers, W. G. W. M. E., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*(3), 403–410.

Anderson, L. A., Palmer, T., Price, N. C., Bornemann, S., Boxer, D. H., & Pau, R. N. (1997). Characterisation of the molybdenum-responsive ModE regulatory protein and its binding to the promoter region of the modABCD (molybdenum transport) operon of *Escherichia coli*. *European Journal of Biochemistry, 246*(1), 119–126.

Balan, A., Santacruz-Pérez, C., Moutran, A., Ferreira, L. C. S., Neshich, G., & Barbosa, J. A. R. G. (2008). Crystallographic structure and substrate-binding interactions of the molybdate-binding protein of the phytopathogen *Xanthomonas axonopodis* pv. citri. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, *1784*(2), 393–399.

Bevers, L. E., Hagedoorn, P. L., Krijger, G. C., & Hagen, W. R. (2006). Tungsten transport protein A (WtpA) in pyrococcus furiosus: The first member of a new class of tungstate and molybdate transporters. *Journal of Bacteriology, 188*(18), 6498–6505.

Bevers, L. E., Hagedoorn, P. L., Santamaria-Araujo, J. A., Magalon, A., Hagen, W. R., & Schwarz, G. (2008). Function of MoaB proteins in the biosynthesis of the molybdenum and tungsten cofactors. *Biochemistry, 47*(3), 949–956.

Blasco, F., Dos Santos, J. P., Magalon, A., Frixon, C., Guigliarelli, B., Santini, C. L., & Giordano, G. (1998). NarJ is a specific chaperone required for molybdenum cofactor assembly in nitrate reductase A of *Escherichia coli*. *Molecular Microbiology, 28*(3), 435–447.

Chan, C. S., Howell, J. M., Workentine, M. L., & Turner, R. J. (2006). Twin-arginine translocase may have a role in the chaperone function of NarJ from *Escherichia coli*. *Biochemical and biophysical research communications, 343*(1), 244–251.

Chao, K. J., Wu, C. N., Chang, H., Lee, L. J., & Hu, S. F. (1997). Incorporation of vanadium in mesoporous MCM-41 and microporous AFI zeolites. *The Journal of Physical Chemistry B, 101*(33), 6341–6349.

Dawson, R. J., Hollenstein, K., & Locher, K. P. (2007). Uptake or extrusion: Crystal structures of full ABC transporters suggest a common mechanism. *Molecular Microbiology, 65*(2), 250–257.

Feng, C., Ma, J., Li, H., Zeng, R., Guo, Z., & Liu, H. (2009). Synthesis of molybdenum disulfide (MoS2) for lithium ion battery applications. *Materials Research Bulletin, 44*(9), 1811–1815.

Fenske, D., Gnida, M., Schneider, K., Meyer-Klaucke, W., Schemberg, J., Henschel, V., et al. (2005). A new type of metalloprotein: The Mo storage protein from *Azotobacter vinelandii* contains a polynuclear molybdenum–oxide cluster. *ChemBioChem, 6*(2), 405–413.

Gerber, S., Comellas-Bigler, M., Goetz, B. A., & Locher, K. P. (2008). Structural basis of trans-inhibition in a molybdate/tungstate ABC transporter. *Science, 321*(5886), 246–250.

González-Ballester, D., Camargo, A., & Fernández, E. (2004). Ammonium transporter genes in *Chlamydomonas*: The nitrate-specific regulatory gene Nit_2 is involved in Amt1; 1 expression. *Plant Molecular Biology, 56*(6), 863–878.

Grunden, A. M., & Shanmugam, K. T. (1997). Molybdate transport and regulation in bacteria. *Archives of Microbiology, 168*(5), 345–354.

Hagen, W. R. (2011). Cellular uptake of molybdenum and tungsten. *Coordination Chemistry Reviews, 255*(9), 1117–1128.

Hernandez, J. A., George, S. J., & Rubio, L. M. (2009). Molybdenum trafficking for nitrogen fixation. *Biochemistry, 48*(41), 9711–9721.

Higgins, C. F. (1992). ABC transporters: From microorganisms to man. *Annual Review of Cell Biology, 8*(1), 67–113.

Holland, I. B., & Blight, M. A. (1999). ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *Journal of Molecular Biology, 293*(2), 381–399.

Hollenstein, K., Comellas-Bigler, M., Bevers, L. E., Feiters, M. C., Meyer-Klaucke, W., Hagedoorn, P. L., & Locher, K. P. (2009). Distorted octahedral coordination of tungstate in a subfamily of specific binding proteins. *JBIC Journal of Biological Inorganic Chemistry, 14*(5), 663–672.

Hollenstein, K., Frei, D. C., & Locher, K. P. (2007). Structure of an ABC transporter in complex with its binding protein. *Nature, 446*(7132), 213–216.

Hu, Y., Rech, S., Gunsalus, R. P., & Rees, D. C. (1997). Crystal structure of the molybdate binding protein ModA. *Nature Structural and Molecular Biology, 4*(9), 703–707.

Imperial, J., Hadi, M., & Amy, N. K. (1998). Molybdate binding by ModA, the periplasmic component of the *Escherichia coli* mod molybdate transport system. *Biochimica et Biophysica Acta (BBA)-Biomembranes, 1370*(2), 337–346.

Lawson, D. M., Williams, C. E., Mitchenall, L. A., & Pau, R. N. (1998). Ligand size is a major determinant of specificity in periplasmic oxyanion-binding proteins: The 1.2 Å resolution crystal structure of *Azotobacter vinelandii* ModA. *Structure, 6*(12), 1529–1539.

Llamas, A., Kalakoutskii, K., & Fernández, E. (2000). *Plant, Cell and Environment, 23*, 1247–1255.

Maupin-Furlow, J. A., Rosentel, J. K., Lee, J. H., Deppenmeier, U., Gunsalus, R. P., & Shanmugam, K. T. (1995). Genetic analysis of the modABCD (molybdate transport) operon of *Escherichia coli*. *Journal of Bacteriology, 177*(17), 4851–4856.

McNicholas, P. M., Rech, S. A., & Gunsalus, R. P. (1997). Characterization of the ModE DNA-binding sites in the control regions of modABCD and moaABCDE of *Escherichia coli*. *Molecular Microbiology, 23*(3), 515–524.

Mouncey, N. J., Mitchenall, L. A., & Pau, R. N. (1995). Mutational analysis of genes of the mod locus involved in molybdenum transport, homeostasis, and processing in *Azotobacter vinelandii*. *Journal of Bacteriology, 177*(18), 5294–5302.

Nichols, J., & Rajagopalan, K. V. (2002). *Escherichia coli* MoeA and MogA function in metal incorporation step of molybdenum cofactor biosynthesis. *Journal of Biological Chemistry, 277*(28), 24995–25000.

Nichols, J. D., Xiang, S., Schindelin, H., & Rajagopalan, K. V. (2007). Mutational analysis of *Escherichia coli* MoeA: Two functional activities map to the active site cleft. *Biochemistry, 46*(1), 78–86.

Rech, S., Deppenmeier, U., & Gunsalus, R. P. (1995). Regulation of the molybdate transport operon, modABCD, of *Escherichia coli* in response to molybdate availability. *Journal of Bacteriology, 177*(4), 1023–1029.

Rech, S., Wolin, C., & Gunsalus, R. P. (1996). Properties of the periplasmic ModA molybdate-binding protein of *Escherichia coli*. *Journal of Biological Chemistry, 271*(5), 2557–2562.

Rosentel, J. K., Healy, F., Maupin-Furlow, J. A., Lee, J. H., & Shanmugam, K. T. (1995). Molybdate and regulation of mod (molybdate transport), *fdhF*, and *hyc* (formate hydrogenlyase) operons in *Escherichia coli*. *Journal of Bacteriology, 177*(17), 4857–4864.

Schwarz, G., Hagedoorn, P. L., & Fischer, K. (2007). Molybdate and tungstate: Uptake, homeostasis, cofactors, and enzymes. In *Molecular microbiology of heavy metals* (pp. 421–451). Berlin: Springer.

Self, W. T., Grunden, A. M., Hasona, A., & Shanmugam, K. T. (2001). Molybdate transport. *Research in Microbiology, 152*(3), 311–321.

Tejada-Jiménez, M., Llamas, Á., Sanz-Luque, E., Galván, A., & Fernández, E. (2007). A high-affinity molybdate transporter in eukaryotes. *Proceedings of the National Academy of Sciences, 104*(50), 20126–20130.

Walkenhorst, H. M., Hemschemeier, S. K., & Eichenlaub, R. (1995). Molecular analysis of the molybdate uptake operon, *modABCD*, of *Escherichia coli* and *modR*.

Wang, R., Okamoto, M., Xing, X., & Crawford, N. M. (2003). Microarray analysis of the nitrate response in *Arabidopsis* roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. *Plant Physiology, 132*(2), 556–567.

Chapter 4
Molybdenum Cofactor-Containing Enzymes and Pathways

Abstract To date, more than 50 enzymes are known to contain Mo and most of them occur in prokaryotes while only six were found in eukaryotes. In all organisms studied so far, the Mo-cofactor is synthesized by a highly conserved biosynthetic pathway that can be divided into four steps, each producing a specific biochemical intermediate. Different nomenclatures were introduced for genes and gene products involved in Mo-co formation according to the organism they belong. On the basis of the active site structure and catalytic function, molybdenum-dependent enzymes can be grouped into two categories: bacterial nitrogenases and pterin-based enzymes. Human and plants Mo-co deficiency are also described.

All enzymes whose function depends on molybdenum catalyse redox reactions by taking advantage of the versatile redox chemistry of this metal, which is controlled by the cofactor itself and the enzyme's environment (Hille 2002). Within these enzymes, molybdenum shuttles between three oxidation states (+4, +5 and +6), thereby catalysing two-electron redox reactions. In most cases, regeneration of the active site involves single-electron steps, resulting in a paramagnetic molybdenum intermediate. To date, more than 50 enzymes are known to contain Mo and most of them occur in prokaryotes while only six were found in eukaryotes (Sigel and Sigel 2002; Magalon et al. 2011; Schwarz et al. 2009).

Molybdenum enzymes are found in nearly all organisms, with *Saccharomyces* as a sharp eukaryotic exception (Zhang and Gladyshev 2008). It is also known that many anaerobic Archaea and some Bacteria are molybdenum-independent but require tungsten for their growth. Tungstate, which is 100-fold less abundant than molybdate, is enriched in deep-sea hydrothermal vents, reflecting conditions on the primitive Earth. Accordingly, many of the known tungsten-dependent hyperthermophilic Bacteria and Archaea are found in such vents (Bevers et al. 2008).

L. Presta et al., *Molybdenum Cofactors and Their role in the Evolution of Metabolic Pathways*, SpringerBriefs in Biometals, DOI 10.1007/978-94-017-9972-0_4

4.1 Molybdenum Cofactor Biosynthesis

In all organisms studied so far, the Mo-cofactor is synthesized by a highly conserved biosynthetic pathway that can be divided into four steps according to its biosynthetic intermediates: cyclic pyranopterin monophosphate (cPMP, previously also known as precursor Z), molybdopterin or metal-containing pterin MPT, adenylated MPT and, at last, Mo-cofactor. In plants (Mendel and Schwarz 2002), fungi (Millar et al. 2001), and humans (Reiss et al. 1998; Stallmeyer et al. 1999a, b), the products of six gene catalyzing Mo-co biosynthesis have been identified. These genes are homologous to their bacterial counterparts and some of the eukaryotic Mo-co biosynthesis genes

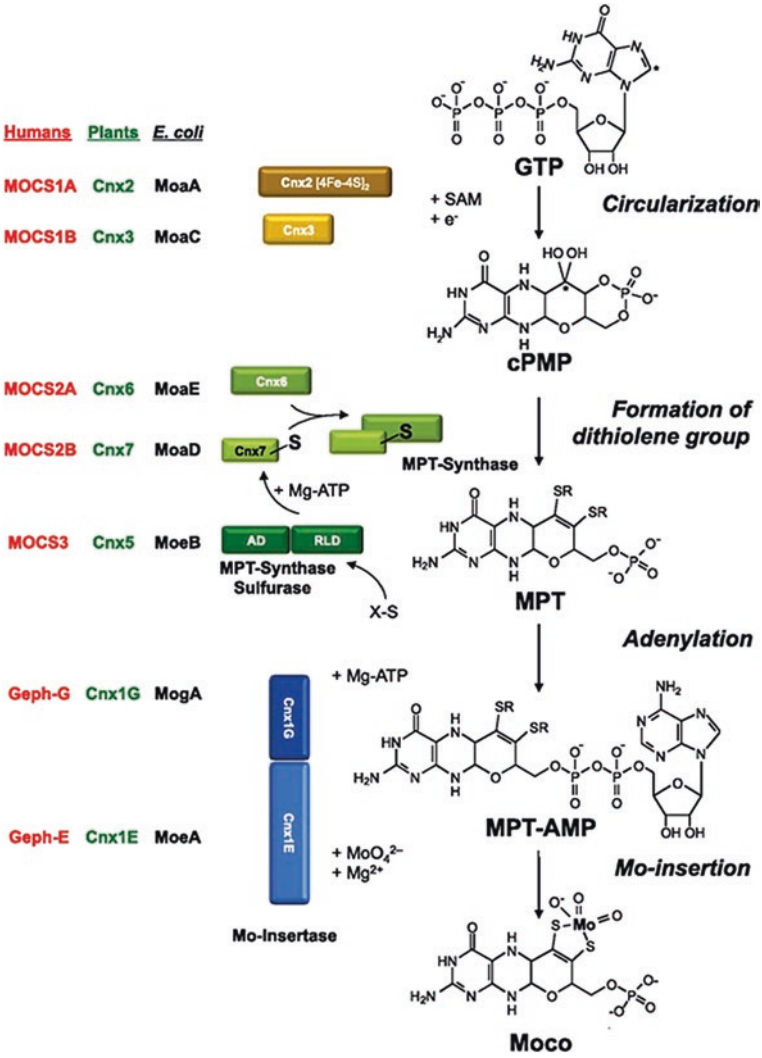

Fig. 4.1 Biosynthesis of eukaryotic Moco. The biosynthesis pathway is divided into four steps, as shown in *italics* on the *right*. On the *left*, the names for the proteins from humans (*red*), plants (*green*), and *E. coli* (*black*) catalyzing the respective steps are given. For MPT and MPT-AMP, the ligands of the dithiolate sulfurs are indicated by *R*, as it is currently unknown at which step copper is bound to the dithiolate. In GTP, the C8 atom of the purine is labeled with an *asterisk*. This carbon is inserted between the 2′- and 3′-ribose carbon atoms, thus forming the new C1′ position in the four-carbon side chain of the pterin (labeled with an *asterisk* in cPMP). In step 2, the heterotetrameric MPT synthase complex converts cPMP into MPT. In this process, two sulfur atoms need to be transferred from the thiocarboxylated C termini of the small subunits, which later form the dithiolene group of MPT. Having transferred their sulfur atoms, the small subunits need to be reloaded with sulfur, which is facilitated by the MPT synthase sulfurase. This enzyme consists of two domains, with the N-terminal adenylation domain (*AD*) catalyzing the Mg-ATP-dependent adenylation at the C-terminal carboxyl group of the small subunit of MPT synthase and the C-terminal RLD being responsible for subsequent sulfur transfer. In step 3, the two-domain protein molybdenum insertase catalyzes the Mg-ATP-dependent adenylation of MPT at its G-domain, with the subsequent transfer of MPT-AMP to its E-domain. In step 4 and occurring at the E-domain, MPT-AMP is deadenylated, and the molybdate anion is incorporated to form mature Moco. *Geph-G* and *Geph-E*, gephyrin G- and E-domains. Reproduced from Mendel (2013). With kind permission of © Journal of Biological Chemistry

are able to functionally complement the matching bacterial mutants. Different nomenclatures were introduced for genes and gene products involved in Mo-co formation according to the organism they belong. In plants genes and their encoded proteins were named according to the *cnx* nomenclature (cofactor for nitrate-reductase and xanthine-dehydrogenase). For human Mo-co genes was introduced a different MOCS (Molybdenum cofactor synthesis) nomenclature. A third nomenclature is used in bacteria. Each step of the Mo-co biosynthesis with the relative nomenclature of involved genes in different organism is shown in Fig. 4.1.

4.1.1 Conversion of GTP into cPMP

In the first step of Mo-co biosynthesis 5′-GTP is transformed into a sulphur-free pterin compound, the cyclic pyranopterin monophosphate (or cPMP), which already possesses the Mo-co-typical four- carbon side chain. In comparison to other intermediate compounds, cPMP is the most stable one with an estimated half-life of several hours at a low pH (Wuebbens and Rajagopalan 1993). Recently, mass spectrometry and 1H NMR analyses revealed that cPMP already possesses a fully reduced tetrahydropyranopterin structure, although still sulfur-free, and is predominantly hydrated at the C1′ position resulting in a geminal diol (Santamaria-Araujo et al. 2004). GTP labeling studies and NMR demonstrated that each carbon atom of the ribose and of the guanine ring is incorporated into cPMP (Wuebbens and Rajagopalan 1995; Rieder et al. 1998). The detailed mechanism of this reaction step remains unclear, yet hypothetical multistep- reactions have been suggested (Wuebbens and Rajagopalan 1995; Rieder et al. 1998). In *A. thaliana* cPMP synthesis is catalyzed by the proteins Cnx2 and Cnx3, encoded by *cnx2* and *cnx3* genes, which were identified using

the approach of functional complementation of *E. coli* Mo-co mutants *moaA* and *moaC*, thus demonstrating the functional conservation between bacterial and plant Mo-co synthesis (Hoff et al. 1995). Cnx2 belongs to the superfamily of S-adenosylmethionine(SAM)-dependent radical enzymes (Hänzelmann et al. 2004). Members of this large family catalyze the formation of protein and/or substrate radicals by reductive cleavage of SAM by a (4Fe–4S) cluster (Sofia et al. 2001). Cnx2 and Cnx3 are larger than their bacterial counterparts because of the N-terminal extensions carrying targeting motifs for mitochondrial or chloroplast transport (Hoff et al. 1995). In fact, both of them have been localized in the matrix of mitochondria where their presence is reasonable, from a physiological point of view, because the mitochondrial matrix is equipped to provide (i) 5′-GTP as substrate for cPMP synthesis, (ii) Fe–S clusters as the essential prosthetic group for Cnx2 (Balk and Lobréaux 2005) and (iii) a reducing environment for stabilization of these oxygen-sensitive clusters bound to Cnx2. The fact that Cnx2 is likely to have (4Fe–4S) clusters similar to its human homolog MOCS1A (Hänzelmann et al. 2004) might explain the mitochondrial localization as this organelle is a major site of plant Fe–S-cluster synthesis (Lill and Mühlenhoff 2006). Therefore, since all the other steps of Mo-co biosynthesis in Eukaryotes occur in the cytoplasm, cPMP formed in this step needs to be transported outside the organelle. Despite the molecule is hydrophobic enough to cross the mithocondrial membrane simply by diffusion, several transport systems in different organisms have been invoked to explain its presence in the cytoplasm. In agreement with this hypothesis a recent work on plants demonstrated that a specific transporter in the inner membrane of mitochondria is involved in the transport of cPMP into the cytosol (Teschner et al. 2010). Therefore, it became of interest to study where the remaining steps of Mo-co biosynthesis take place. From another point of view, this question became imperative as several biosynthetic intermediates, such as cPMP and MPT, are labile in solution and need protection. The existence of a hypothetical Mo-co biosynthetic multi-enzyme complex was discussed in the past by Mendel and Bittner (2006) according to which this complex would facilitate product–substrate flow, resulting in micro-compartmentalization and ensuring a swift and protected transfer of the labile intermediates within the reaction sequence from GTP to Mo-co.

4.1.2 Synthesis of Molybdopterin

In the second stage sulfur is transferred to cPMP in order to generate molybdopterin, also called metal-containing pterin (MPT) since W can be found in the active site instead of Mo. This reaction is catalyzed by the enzyme MPT synthase, a heterotetrameric complex of two small and two large subunits that stechiometrically converts cPMP into MPT. Plant MPT-synthase is encoded by *cnx6* and *cnx7* genes, encoding the large and the small subunits, respectively. The sulfur is bound to the C-terminus of the small subunits as thiocarboxylate.

Due to the fact that each small subunit of MPT synthase carries a single sulfur atom, a two-step mechanism for the formation of the MPT dithiolate has been proposed, which involves the formation of a mono-sulfurated intermediate (Gutzke et al. 2001; Wuebbens and Rajagopalan 1993). After MPT-synthase has transferred the two sulfurs to cPMP, it has to be re-sulfurated by the MPT-synthase sulfurase in order to reactivate the enzyme for the next reaction cycle of cPMP conversion. This resulfuration is catalyzed by Cnx5 involving very likely an adenylation of MPT-synthase followed by sulfur transfer similar to its human homolog MOCS3 (Matthies et al. 2004, 2005). Cnx5 and MOCS3 are two-domain proteins consisting of a N-terminal domain responsible for adenylating MPT-synthase and a C-terminal rhodanese-like domain where the sulfur is bound to a conserved cysteine in the form of persulfide (Matthies et al. 2005). The identity of the donor for the reactive mobile sulfur is as yet unknown, but the contribution of a persulfide-generating cysteine desulfurase is possible (Leimkühler et al. 2001).

4.1.3 Adenlyation of Molybdopterin

After synthesis of the MPT moiety, the chemical backbone is built for binding and coordination of the Mo atom. In the third step, therefore, Mo has to be transferred to MPT in order to form Moco, thus linking the molybdate uptake system to the MPT pathway. Mutants defective in this step produce MPT and can be partially repaired by growing them on high-molybdate medium. In bacteria, this step is catalyzed by two proteins which during evolution to higher organisms were fused to a two-domain protein. Earlier it was assumed that one domain should be essential for generating an activated form of Mo that is incorporated by the other domain into bound MPT (Schwarz et al. 1997a, b, 2000). Recently the exact mechanism was uncovered in plants where the protein Cnx1 is catalyzing this step (Llamas et al. 2004). The C-terminal Cnx1 domain (Cnx1-G) was known to tightly bind MPT (Schwarz et al. 1997a, b). Yet, its crystal structure (Kuper et al. 2004) revealed an unexpected finding: a novel reaction intermediate, adenylated MPT (MPT-AMP). Subsequently it was demonstrated that Cnx1-G adenylates MPT in an Mg2+- and ATP-dependent way and forms MPT-AMP that remains bound to Cnx1-G (Llamas et al. 2004).

The crystal structure of the Cnx1-G revealed another unexpected finding, namely a copper bound to the MPT dithiolate sulfurs, whose nature was confirmed by anomalous scattering of the metal. In both structures the copper atom shows tetragonal coordination with two waters as additional ligands in the MPT-bound state, while one of these waters is replaced by a histidine in the MPT-AMP-bound structure. Up to now the function of this novel MPT ligand is unknown but copper might play a role in sulfur transfer to cPMP, in protecting the MPT dithiolate from oxidation, and/or presenting a suitable leaving group for Mo insertion. The origin of this copper is still unclear but it is reasonable to assume that it binds to the enedithiolate group just after the latter has been formed, i.e. at the end of step 2 of Mo-co biosynthesis.

4.1.4 Mo-Co Assembly

In the fourth and final step, mature Mo-co is formed by the ligation of a single Mo atom to adenylated molybdopterin; this process directly depends on the intracellular availability of molybdenum. MPT-AMP is transferred to the N-terminal domain of Cnx1 (Cnx1-E) thereby building a product–substrate channel. Cnx1-E cleaves the adenylate, releases copper, and inserts Mo, thus yielding active Mo-co. This reaction was coupled to the metal exchange reaction where bound copper was released and Mo was transferred to MPT thus yielding mature Mo-co. As copper is always found in a protein-bound state it might be that Cnx1 interacts with a copper chaperone when the metal is released during Mo insertion. Using a split-ubiquitin based two-hybrid approach with Cnx1 as bait, a copper chaperone homolog has been recently identified (Mendel and Bittner 2006). In vitro studies with Cnx1-G-bound MPT-AMP revealed an inhibition of Mo-co synthesis in the presence of 1 mM CuCl2, providing a link between Mo and copper metabolism (Kuper et al. 2004). Copper inhibition of Mo-co synthesis can be explained by inhibition of the Mg-dependent Mo insertion reaction. The latter is supported by the suppression of copper inhibition with equimolar amounts of Cnx1-E and is in line with the known copper inhibition of pyrophosphatases (Moorhead et al. 2003). Mo-co deficiency might occur when cellular copper concentrations are increased, as seen in human patients affected with Wilson's disease (Mercer 2001), where copper accumulates in liver and brain, resulting in severe damage to both organs. In plants, one should consider that the toxicity of elevated copper levels could also go back—at least in part—to a distorted Mo-co biosynthesis. However, it is to be noted that copper shortage also should be detrimental for Mo-co biosynthesis. Therefore, when studying copper homeostasis the accompanying analysis of Mo metabolism could shed further light onto the link between Mo and copper homeostasis.

4.1.5 Further Biochemical Modifications

The complete Mo-co appears in four basic configurations, whose structures have been determined by a combination of rigorous biochemical studies and confirmed by the emergence of a large amount of structural data (Johnson and Rajagopalan 1982; Johnson et al. 1990, 1991; Hilton and Rajagopalan 1996; Solomon et al. 1997). The cofactor typically has a tricyclic pyranopterin structure, with pyrimidine, pyrazine, and pyran rings, respectively. The alternative form of the pterin is the bicyclic molybdopterin form, which contains the pyrimidine and pyrazine rings but contains an open pyran ring. Mo and W coordination is via two dithiolene sulfurs, which are attached to the pyran ring. Also attached to this ring is a phosphomethyl group, and it is the presence of a terminal phosphate that has led to this form of the cofactor being referred to as a mononucleotide. Given the dominance of the pyranopterin structure (tricyclic) over the original molybdopterin (bicyclic) structure developed on the basis of chemical studies

(Johnson and Rajagopalan 1982; Johnson et al. 1990, 1991), it is referred to as Mo-PPT or W-PPT when complexed with Mo or W, respectively. Mo-PPT is the form of the cofactor found in eukaryotes and in essentially all sulfite oxidase enzymes (SUOX) and some xanthine dehydrogenase enzymes (XDH). PPT can be modified by the addition of a nucleotide (typically cytosine or guanine) (Rajagopalan and Johnson 1992) at the phosphomethyl position to generate a dinucleotide form. Intriguingly, two PPTs can coordinate a single W atom via a pair of dithiolenes. This form of the cofactor is the bis-mononucleotide form. The vast majority of bacterial molybdoenzymes contain the molybdo-bis (pyranopterin guanine dinucleotide) (Mo-bisPGD) cofactor or its tungsto derivative (W-bisPGD). A minority of the Mo-bisPGD cofactors contain both a tricyclic pyranopterin and a bicyclic molybdopterin with a structure similar to that originally proposed for the cofactor by Rajagopalan and co-workers (Bertero et al. 2003; Johnson and Rajagopalan 1982; Kloer et al. 2006).

Cofactor nomenclature is further complicated in those enzyme subunits that contain both a Mo-bisPGD and an (4Fe–4S) cluster that is at the terminus of an electron transfer relay that can comprise up to five (Fe–S) clusters and two hemes (Rothery et al. 2008). In the original structure determinations of the Mo-bisPGD containing dimethyl sulfoxide reductases (DorA), the two pterins were arbitrarily referred to as “P” and “Q” (Schindelin et al. 1996; Scheneider et al. 1996). When these subunits are structurally aligned to those that also contain an (Fe–S) cluster (e.g. FdhF from *E. coli*) (Boyington et al. 1997), the P-pterin is coincidentally located proximal to the (4Fe–4S) cluster, whereas the Q-pterin is located distal to it hence it can be referred to also as the D-pterin, or distal pterin.

4.1.6 Regulation of Mo-Co Biosynthesis in Bacteria

Genetic analysis of the regulation of Mo-co biosynthesis demonstrated that the *moaABCDE* genes are operonically organized (Rivers et al. 1993) and that this is the main target for transcriptional and translational regulation of Mo-co biosynthesis (Anderson et al. 2000). Expression of this operon is enhanced under anaerobic conditions and in the presence of molybdate. The *E. coli* dimeric ModE protein binds two molybdate ions and this complex acts as a positive regulator that binds to the *moa* promotor region (Fig. 4.2) and enhances transcription of the operon; moreover, the same ModE/molybdate complex has a positive effect on molybdoenzymes transcription, e.g. on the operons encoding DMSO reductase and nitrate reductase A (McNicholas et al. 1998).

As described above (see Fig. 3.2), it acts as a transcriptional repressor of the *modABCD* molybdate transport operon when molybdate is already present in sufficient quantities in the cell (Self et al. 1999; Schüttelkopf et al. 2003). In addition to the ModE/molybdate complex, transcription of the *E. coli moaABCDE* operon is also up-regulated by the FNR protein (fumarate and nitrate reduction regulatory protein) (Anderson et al. 2000). FNR is a transcriptional regulator that is essential

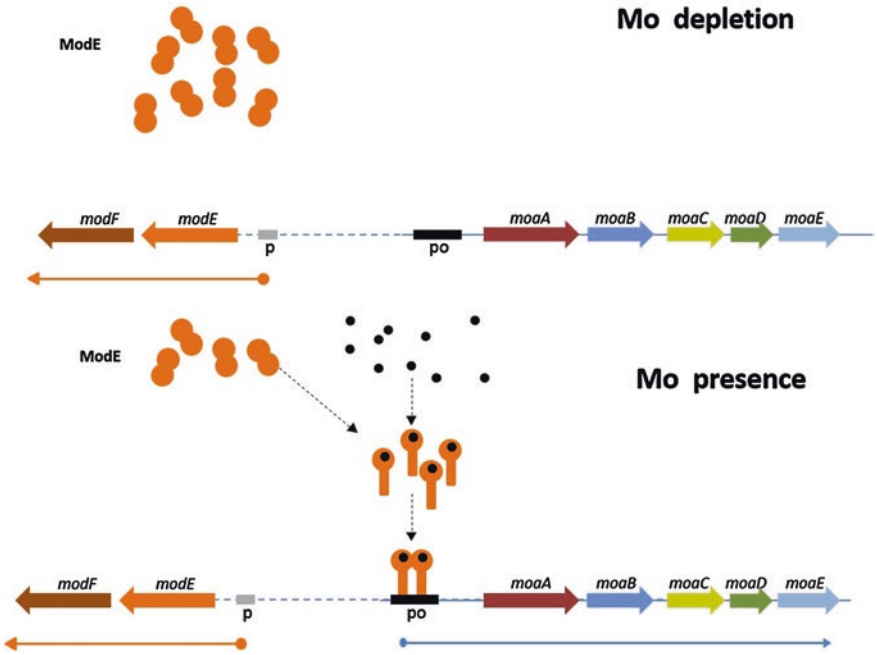

Fig. 4.2 Regulation of moaABCDE operon by ModE-molybdenum complex

for expressing the proteins involved in anaerobic respiratory processes (Unden et al. 2002). FNR directly senses the ambient oxygen concentration via the disassembly and reassembly of its (4Fe–4S) clusters (Unden and Bongaerts 1997). The combined regulation of *moaABCDE* by both ModE and FNR ensures that transcription of the Mo-co biosynthesis proteins only occurs under conditions where they are both needed (anaerobic growth conditions as sensed by FNR) and capable of producing active Mo-co (sufficient molybdate levels as sensed by ModE) (Anderson et al. 2000).

Another gene cluster involved in the biosynthesis of Mo-co, is the bicistronic *moeAB* operon. It encodes two proteins, MoeA and MoeB, but only the physiological role of the second protein is known. The MoeB protein activates MPT synthase, which catalyzes the conversion of MPT precursor to MPT by introducing the needed sulfur to which Mo is coordinated in the molybdenum cofactor. Since the MoeB protein acts as an S donor in MPT synthesis, it is possible that the product of the first gene, MoeA protein, also has a similar role in linking S metabolism and Mo metabolism in the cell (Hasona et al. 1998). Regulation of the *moeAB* operon is more complicated, since it is regulated by FNR, ArcA (the DNA-binding response regulator in a two-component regulatory system in which ArcB represses transcription of aerobic genes and activates some anaerobic genes) and NarL (part of a two-component regulator protein system for nitrate/nitrite response that acts with the sensor protein NarX) (Hasona et al. 2001). Surprisingly, FNR represses the nitrate-dependent transcription of the moe

operon, but it is assumed that under anaerobic conditions, an intermediate level of transcription from the moeAB operon is ensured by the antagonistic effects of FNR and ArcA-P. In addition, NarL regulates a nitrate-dependent increase in *moeAB* expression (Hasona et al. 2001).

The observation that transcription of the *moaABCDE* operon is decreased when Mo-co is present in excess implies that there is an additional level of regulation in the system that is not mediated by FNR or ModE (Anderson et al. 2000). The possible existence of a regulatory region located upstream of the *moaA* start codon that was responsive to the availability of Moco was first described by Anderson et al. (2000). Its existence was confirmed in 2008 when Regulski et al. (2008) identified a highly conserved RNA motif upstream of the *moa* operon in *E. Coli*, which was also present upstream of genes encoding molybdoenzymes in a number of different organisms. They determined that the RNA molecule encoded by this sequence controlled gene expression in response to Mo-co production and suggested that this family of RNA sequences represented a novel class of riboswitches that sense Mo-co (Regulski et al. 2008). Riboswitches are structured RNA domains that selectively bind metabolites or metal ions and function as gene control elements (Pramanik et al. 2004). Prior to the identification of the Mo-co RNA family, four previously identified classes of riboswitch aptamers were known to exist in *E. coli*. These respond to coenzyme B12, thiamine pyrophosphate, flavin mononucleotide and lysine (Pramanik et al. 2004).

Regulski et al. (2008) determined that the architectural features of the Mo-co RNA, which spans 138 nucleotides and forms at least five conserved stem-loop elements, are similar to other classes of riboswitches. They found that translation of the *moa* operon is prevented when the Mo-co RNA structure and Mo-co are present, although a direct binding of Mo-co to this RNA motif has not been demonstrated to date due to the lability of isolated Mo-co (Regulski et al. 2008). They also suggested that only the molybdenum derivative of Mo-co binds to the RNA region and not the tungsten derivative, which would have its own class of riboswitches. This would imply that Mo-co RNAs are able to distinguish between these two nearly identical cofactors (Regulski et al. 2008). A Mo-co sensing riboswitch would ensure that translation of the proteins encoded by the *moaABCDE* operon is down regulated when Mo-co is already present in sufficient amounts. Thus riboswitch regulation, in combination with ModE and FNR regulation, would allow for a rapid and efficient response to the changing demand for Mo-co within the cell.

4.2 Molybdenum Containing-Enzymes

On the basis of the active site structure and catalytic function, molybdenum-dependent enzymes can be grouped into two categories:

1. bacterial nitrogenases containing a FeMo-co in the active site (discussed in Chap. 5);
2. pterin-based (MoCo) molybdenum enzymes.

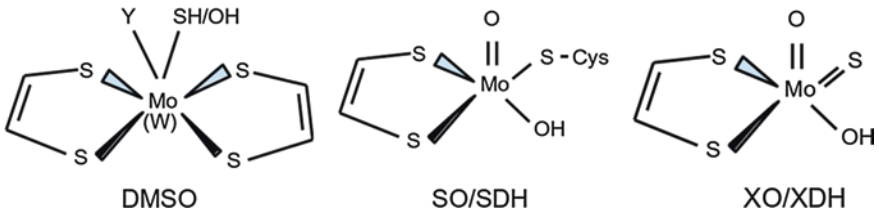

Fig. 4.3 Classification of Mo and W pyranopterin-dependent enzymes. Abbreviations: DMSO or dimethyl sulphoxide reductase family, SO/SDH sulfite oxidase and sulphite dehydrogenase reductase family XO/XDH xanthine oxidase and xanthine dehydrogenase family

With the only exception of the multinuclear MoFe7 cluster present in nitrogenase, molybdenum is found in all other known molybdoenzymes in a mononuclear form, which possesses an organic tricyclic pyranopterin cofactor coordinated to the metal, as described above (Romão 2009). In literature the cofactor is often called molybdopterin or Mo-co, since it was originally believed to be present only in molybdenum enzymes. However, it was later found that the same form is present in tungsten enzymes; hence, it is possible to add a third family to those previously listed, the so-called "true tungstoenzymes" to which belongs the archeal aldehyde ferrodoxin oxidoreductase (AOR).

Among Bacteria Mo-co–containing enzymes may be generally divided into three categories (Magalon et al. 2011; Schwarz et al. 2009) that are characterized by the types of reactions being catalyzed and are divided into families, exemplified by xanthine oxidase (XO), sulphite oxidase (SO) and dimethyl sulphoxide reductase (DMSOR), which each have a distinct active site structure. A scheme with the classification of Mo and W pyranopterin-dependent enzymes is shown in Fig. 4.3. Mo-enzymes are required for diverse key reactions in the global carbon, sulfur, and nitrogen metabolism, and although many Mo-enzymes have been found in the three domains of life, their number is still increasing. As mentioned before, only a limited number of Mo-enzymes is present in eukaryotes, where they can be subdivided into two families: the xanthine oxidase (XO) family represented by xanthine dehydrogenase (XDH), aldehyde oxidase (AO), pyridoxal oxidase and nicotinate hydroxylase, and the sulfite oxidase (SO) family represented by sulfite oxidase (SO) and nitrate reductase (NR). The mitochondrial amidoxime reducing component (mARC), which has been identified recently in pig liver mitochondria (Havemeyer et al. 2006), has not yet been integrated into one of the aforementioned families and at present its formation of a new family of Mo-enzymes cannot be excluded.

4.2.1 XO Family

The members of this group are molybdo-flavoenzymes that catalyze the oxidative hydroxylation of a wide range of aldehydes and aromatic heterocycles. These enzymes require a final step of maturation prior to or after insertion of Mo-co.

In addition to the dithiolene sulfurs of the pterin moiety and two oxo groups, the molybdenum atom of Mo-co needs the addition of a terminal inorganic sulfur to provide enzyme activity. This final step is catalyzed by the Mo-co sulfurase protein (ABA3 in plants and HMCS in humans) (Mendel 2013).

- XDH is a key enzyme of purine degradation and oxidizes hypoxanthine to xanthine and xanthine to uric acid by simultaneous release of electrons from the substrate. The enzymes of this family are involved in two-electron transfer hydroxylation and oxo-transfer reactions with water as the source of oxygen (Romão et al. 1995). The Mo active site of these enzymes is characterized by a cyanolysable sulfur ligand, typically a sulfido group. Mo sulfuration, and thus enzyme activation, occurs either concurrently or after Mo-PPT/PCD insertion into the enzyme (Hille 1996; Mendel and Bittner 2006; Neumann et al. 2009). In Eukaryotes the enzyme is active as a homodimer composed of two identical subunits of approximately 150 kDa, each being subdivided into three distinct domains: an N-terminal domain with a size of 20 kDa for binding of two (2Fe–2S) clusters, a 40 kDa domain harboring a FAD-binding site, and a C-terminal domain required for Moco-binding and dimerization. Electrons derived from substrate hydroxylation at the Mo-center go via Fe–SII and Fe–SI to the flavin cofactor. At the FAD site, the electrons are transferred either to NAD+ to form NADH, or they are transferred to molecular oxygen to yield superoxide anions. XDH enzymes are discussed not only to decompose purines but also to have additional physiological functions in ROS (reactive oxygen species) metabolism (Mendel and Bittner 2006). Human XDH is discussed to be a pathogenic agent in many forms of ischaemia–reperfusion injury and might also be involved in inflammatory signal transduction (Harrison 2002; Meneshian and Bulkley 2002). The ROS production of plant XDH might also be of physiological importance because increasing XDH activities and simultaneous ROS production were observed upon plant–pathogen interactions (Montalbini 1992), hypersensitive response (Montalbini et al. Montalbini and Della Torre 1996), drought stress (Yesbergenova et al. 2005) and natural senescence (Pastori and del Rio 1997; Hesberg et al. 2004). Prokaryotic XDH's are more complicated in structure. For example, Pseudomonas putida Quinoline 2-oxidoreductase (QorMLS) contains three separate subunits for each type of cofactor: i.e. a FAD binding subunit (QorM), Mo-PCD binding subunit (QorL), and QorS which binds two non-identical (2Fe–2S) clusters (Bonin et al. 2004; Bläse et al. 1996). Alternatively, Rhodobacter capsulatus XDH is a dimer of dimers combining a fusion of the FAD and (2Fe–2S) subunits into one. Electron transfer proceeds from Mo → (Fe–S) I → (Fe–S) II → FAD and finally to NAD+. In a different variation Desulfovibrio gigas aldehyde oxidase (DgMOP) does not contain FAD and electrons are transferred through a complex of several subunits with 11 redox center to form hydrogen (Truglio et al. 2002; Romão et al. 1995; Leimkühler et al. 1998; Dietzel et al. 2009; Barata et al. 1993).
- **AO** proteins are cytoplasmic enzymes that catalyze the oxidation of a variety of aromatic and nonaromatic heterocycles and aldehydes, thereby converting them to

the respective carboxylic acid. AO enzymes are very similar to XDH enzymes as they share a high degree of sequence homology, have nearly identical molecular mass, bind the same cofactors, form dimers, and also act as hydroxylases (Mendel and Bittner 2006). Phylogenetic analyses have shown that AO proteins have derived from XDH after ancient gene duplications (Rodríguez-Trelles et al. 2003). Differently from XDH, AO enzymes are strict oxidases that are unable to bind NAD+ and exclusively use molecular oxygen as electron acceptor. In Eucaryotes this class of enzymes is known to be involved in the biosynthesis of several hormones as vitamin A in humans and abscisic acid and/or auxins in plants.

4.2.2 SO Family

Members of this family include not only sulfite oxidizing enzymes found in plants, animals and bacteria but also the assimilatory eukaryotic nitrate reductases. These assimilatory nitrate reductases catalyze the first and rate-limiting steps of nitrate assimilation in plants, algae and fungi, and are totally different from the bacterial nitrate reductases with respect to primary and tertiary structures (Romão 2009) the latter, in fact, belong to the DMSO reductase family of molybdenum enzymes and are considered in the next section.

There are three different sulfite-oxidizing enzymes presently known: sulfite oxidase (SO) in animals, SO in plants, and sulphite dehydrogenase (SDH) in bacteria.

SO catalyzes the physiologically-vital oxidation of sulfite to sulphate, a two-electron transfer reaction in which the electrons from sulfite reduce the Mo-center from Mo VI to Mo IV. In animals electrons are transferred to the heme Fe III of the cytochromeb5 domain and finally from FeII to cytochrome c, the final physiological electron acceptor, while in plants the enzyme is lacking the cytochrome b5 domain hence the electrons are directly transferred to molecular oxygen with simultaneous formation of hydrogen peroxide. Thus, plant SO is the simplest Mo-enzyme found in eukaryotes with its intramolecular redox center only consisting of Moco.

Still, in plants the physiological role of SO is as yet not absolutely clear while in animals it is well know that this enzyme is critical in the oxidative degradation of the sulfur-containing amino acids cysteine and methionine and it is required for removing sulfite excess from the cell.

Human SO deficiency is a fatal genetic disorder that leads to neurotoxicity and early death.

The bacterial enzyme SDH has an important role in converting sulfite formed during dissimilatory oxidation of reduced sulfur compounds. Enzymes within this family are typically heterodimerics (SorAB) with one molybdopterin cofactor (MPT) and a Mo VI OOH nucleus.

They contain Moco and one c-type heme in independent subunits. Since they cannot transfer electrons to molecular oxygen intraprotein interdomain electron transfer occurs between the two centres and is critical for the enzymatic turnover (Romão 2009).

4.2.3 DMSO Reductase Family

Members of the DMSOR family catalyze the transfer of an oxygen atom to or from a lone electron pair of the substrate and they are found in both bacteria and Archaea. The DMSO reductase family has been divided into subfamilies I, II and III since it includes diverse enzymes such as dissimilatory nitrate reductases and formate dehydrogenases (subfamily I); respiratory nitrate reductases I and ethylbenzene dehydrogenase (EBDH) (subfamily II); DMSO reductase and trimethylamine N-oxide (TMAO) reductase (subfamily III) (Romão 2009).

Despite this heterogeneity inside of the group the crystal structures of enzymes of the DMSOR family show a high degree of similarity in the overall fold. Simple but remarkable local differences at the catalytic sites of the different enzymes are consistent with the stunning diversity of functions performed by them. It is convenient to consider the overall reaction mechanism as consisting of a coupled pair of reductive and oxidative half-reactions, characterized by the reduction of Mo(VI) and the oxidation of Mo(IV), respectively. Because these enzymes catalyze a two-electron redox reaction to or from the substrate, it is likely that the Mo also directly undergoes a two-electron change in oxidation state during the appropriate half-reaction. The return to the resting state of the enzyme must involve either addition or removal of electrons at the Mo, by means of electron transfer between the Mo and a second redox center. Because the second center is typically a one-electron redox group, such as a heme or iron-sulfur cluster, the Mo is restored to the resting state through a sequence of one-electron transfers that will first generate the intermediate Mo(V) state. As an example, the resting state of DMSO reductase is the Mo(IV) state that binds DMSO, followed by reduction of the substrate to produce dimethylsulfide (DMS) and the Mo(VI) state of the enzyme. Two subsequent one-electron transfer reactions regenerate the Mo(IV) state after passing through the Mo(V) state. In most cases, the second redox center involved in these electron transfer reactions is contained within the same enzyme molecule as the Mo-co. One notable exception to this general observation is Rhodobacter's DMSO reductase, which contains no cofactor other than Mo-co. This feature has made this enzyme a valuable target for optical spectroscopy studies (Bastian et al. 1991; Benson et al. 1992) because the spectral features of the Mo-co are dominated by the much larger absorption of the other prosthetic groups, e.g. iron-sulfur clusters, hemes, and flavins.

4.2.4 The Archaeal Aldehyde Oxidoreductase Family

These enzymes catalyze the oxidation of an aldehyde to an acid, with electrons being transferred to a soluble ferredoxin (similar reactions are also catalyzed by a subset of the Family II enzymes, see below). This family is distinguished by the presence of a W-bisPPT cofactor and the best-characterized examples are found in the hyperthermophilic archaeon P. furiosus, with examples

including formaldehyde: ferredoxin oxidoreductase (FOR) (Bevers et al. 2005), glyceraldehyde-3-phosphate:ferredoxin oxidoreductase (GAPOR) (Reher et al. 2007), aldehyde oxidoreductase (AOR) (Bevers et al. 2005), as well as the YdhV subunit of the *E. coli* YdhYVXU operon (Partridge et al. 2008). It is notable that the first structure determined of a mononuclear Mo/W enzyme was that of the P. furiosus AOR (Chan et al. 1995; Roy and Adams 2002). It is a dimer composed of two 605 residue subunits. Each subunit binds a W-bisPPT cofactor and a (4Fe–4S) cluster. The structure revealed several novel features which are applicable to many other Mo-containing enzymes including the finding that the W was coordinated by the dithiolenes of two anti-parallel pyranopterins.

4.3 Human Mo-Co Deficiency

Molybdenum cofactor deficiency in humans is inherited as an autosomal recessive hereditary disorder that results in the pleiotropic loss of SO, XOR and AO activity hence molybdenum is essential in the mechanism of action of these pterin-dependent enzymes which are involved in diseases such as gout, radical damage following cardiac failure and combined oxidase deficiency. The severe phenotype of Moco-deficient patients, although rare, is characterized by progressive neurological damage leading in most cases to early childhood death. In fact symptoms develop shortly after birth when the infant's metabolism starts to operate and toxic metabolites accumulate within the body. Damages are mainly caused by the deficiency of SO that protects the organism, in particular the brain, from elevated levels of toxic sulfite (Johnson 2003). Nevertheless clinical symptoms may result from sulfite toxicity, tissue-specific sulfate deficiency or a combination of both. So far, disease-causing mutations have been identified in three of the four known Mo-co-synthetic human genes: the bicistronic MOCS1, MOCS2 and GEPH (Reiss and Johnson 2003). Given to the new discoveries on the essential genes for human Mo-co biosynthesis, the identification of mutations became possible. Among all known Mo-co-deficient patients studied so far, most patients harbor mutations in the MOCS1 gene resulting in a defect in step 1 of Mo-co biosynthesis, in patients in which this mutation occurs no MPT or active Mo-co are detectable, and consequently all Mo enzyme are affected and their activities are absent (Schwarz 2005).

To date no effective therapy is available, Mo-co deficiency cannot be treated by supplementation with the cofactor because Mo-co is extremely unstable outside the protecting environment of an apoMo enzyme. Among all intermediates of the Mo-co synthetic pathway, cPMP has the highest stability at neutral pH values ($t1/2 = 1$ h) due to its unique pyranopterin nature and the absence of the very reactive dithiolate sulfurs and Mo. Therefore, a method was recently established to produce and purify cPMP in *E. coli* in order to use it for a substitution therapy of Mo-co deficiency in mice. Repeated injections of cPMP into MOCS1-deficient mice resulted in a dose-dependent extension of life span. Given to the treatment MPT levels and Mo enzyme activities were partially restored but stopping cPMP

treatment at any time resulted in a progressive reduction of MPT levels and Mo enzyme activities and death of the animal 10–15 days after receiving the last injection (Schwarz 2005). The first human exposure to cPMP treatment has been reported recently (Schwarz and Veldman 2014), where a patient was diagnosed on day 6 of life, and experimental treatment was started on day 36. Within days, all biomarkers returned to almost normal readings and stayed constant.

4.4 Mo Deficiency in Plants

The common symptoms of Mo deficiency in plants include general yellowing of the leaves, stunted growth, and rolling, curling, and scorching of the leaf margins and, in most cases, they are similar to the symptoms of nitrogen deficiency (Gupta 1997) It was found that Mo deficiency affects plant nitrogen and sulphur metabolisms, in a manner similar (as the nitrate assimilation activity is inhibited by the shortage of Mo-co) but still different from nitrate or sulphate limitation. In a work by Ide et al. (2011), transcriptome analysis has revealed that Mo deficiency in plants had impacts on the genes involved in metabolism, transport, development, stress responses, and signal transduction. The effects are pleiotropic, nevertheless, global alternation of metabolic networks can be captured by integration of the transcript and metabolite profiling approaches. The direct effects of Mo deficiency and the secondary effects of metabolic changes, for example, the increase in sucrose, are indistinguishable. However, the increased accumulation of NR gene transcripts seemed to be one of the Mo-deficiency-responses to compensate for the decrease in Mo-co in both WT and MoT1-, as NR is critical for survival. The molecular mechanisms of plant responses to Mo deficiency is not yet understood, but Mo and the Mo-transporter MOT1 play important roles in maintaining primary metabolism on nitrogen, carbon, and sulphur under Mo deficiency. Further investigation of genes, metabolites, and proteins regulated by Mo is needed to understand the regulatory mechanisms of plant responses to Mo deficiency.

Bibliography

Anderson, L. A., McNairn, E., Leubke, T., Pau, R. N., & Boxer, D. H. (2000). ModE-dependent molybdate regulation of the molybdenum cofactor operon moa in *Escherichia coli*. *Journal of Bacteriology, 182*(24), 7035–7043.

Balk, J., & Lobréaux, S. (2005). Biogenesis of iron–sulfur proteins in plants. *Trends in Plant Science, 10*(7), 324–331.

Barata, B. A., LeGall, J., & Moura, J. J. (1993). Aldehyde oxidoreductase activity in Desulfovibrio gigas: in vitro reconstitution of an electron-transfer chain from aldehydes to the production of molecular hydrogen. *Biochemistry, 32*(43), 11559–11568.

Bastian, N. R., Kay, C. J., Barber, M. J., & Rajagopalan, K. V. (1991). Spectroscopic studies of the molybdenum-containing dimethyl sulfoxide reductase from *Rhodobacter sphaeroides f.* sp. *denitrificans*. *Journal of Biological Chemistry, 266*(1), 45–51.

Benson, N., Farrar, J. A., McEwan, A. G., & Thomson, A. J. (1992). *FEBS Letters, 2*, 169–172.

Bertero, M. G., Rothery, R. A., Palak, M., Hou, C., Lim, D., Blasco, F., et al. (2003). Insights into the respiratory electron transfer pathway from the structure of nitrate reductase A. *Nature Structural & Molecular Biology*, *10*(9), 681–687.

Bevers, L. E., Bol, E., Hagedoorn, P. L., & Hagen, W. R. (2005). WOR5, a novel tungsten-containing aldehyde oxidoreductase from *Pyrococcus furiosus* with a broad substrate specificity. *Journal of Bacteriology, 187*(20), 7056–7061.

Bevers, L. E., Hagedoorn, P. L., Santamaria-Araujo, J. A., Magalon, A., Hagen, W. R., & Schwarz, G. (2008). Function of MoaB proteins in the biosynthesis of the molybdenum and tungsten cofactors. *Biochemistry, 47*(3), 949–956.

Bläse, M., Bruntner, C., Tshisuaka, B., Fetzner, S., & Lingens, F. (1996). Cloning, expression, and sequence analysis of the three genes encoding quinoline 2-oxidoreductase, a molybdenum-containing hydroxylase from *Pseudomonas putida 86*. *Journal of Biological Chemistry, 271*(38), 23068–23079.

Bonin, I., Martins, B. M., Purvanov, V., Fetzner, S., Huber, R., & Dobbek, H. (2004). Active site geometry and substrate recognition of the molybdenum hydroxylase quinoline 2-oxidoreductase. *Structure, 12*(8), 1425–1435.

Boyington, J. C., Gladyshev, V. N., Khangulov, S. V., Stadtman, T. C., & Sun, P. D. (1997). Crystal structure of formate dehydrogenase H: Catalysis involving Mo, molybdopterin, selenocysteine, and an Fe_4S_4 cluster. *Science, 275*(5304), 1305–1308.

Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W., & Rees, D. C. (1995). Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science, 267*(5203), 1463–1469.

Dietzel, U., Kuper, J., Doebbler, J. A., Schulte, A., Truglio, J. J., Leimkühler, S., & Kisker, C. (2009). Mechanism of substrate and inhibitor binding of *rhodobacter capsulatus* xanthine dehydrogenase. *Journal of Biological Chemistry, 284*(13), 8768–8776.

Gupta, U. C. (Ed.). (1997). Molybdenum in agriculture. Cambridge: Cambridge University Press.

Gutzke, G., Fischer, B., Mendel, R. R., & Schwarz, G. (2001). Thiocarboxylation of molybdopterin synthase provides evidence for the mechanism of dithiolene formation in metal-binding pterins. *Journal of Biological Chemistry, 276*(39), 36268–36274.

Hänzelmann, P., Hernández, H. L., Menzel, C., García-Serres, R., Huynh, B. H., Johnson, M. K., et al. (2004). Characterization of MOCS1A, an oxygen-sensitive iron-sulfur protein involved in human molybdenum cofactor biosynthesis. *Journal of Biological Chemistry*, *279*(33), 34721–34732.

Harrison, R. (2002). Structure and function of xanthine oxidoreductase: where are we now? *Free Radical Biology and Medicine, 33*(6), 774–797.

Hasona, A., Ray, R. M., & Shanmugam, K. T. (1998). Physiological and genetic analyses leading to identification of a biochemical role for the *moeA* (molybdate metabolism) gene product in *Escherichia coli*. *Journal of Bacteriology, 180*(6), 1466–1472.

Hasona, A., Self, W. T., & Shanmugam, K. T. (2001). Transcriptional regulation of the *moe* (molybdate metabolism) operon of *Escherichia coli*. *Archives of Microbiology, 175*(3), 178–188.

Havemeyer, A., Bittner, F., Wollers, S., Mendel, R., Kunze, T., & Clement, B. (2006). Identification of the missing component in the mitochondrial benzamidoxime prodrug-converting system as a novel molybdenum enzyme. *Journal of Biological Chemistry, 281*(46), 34796–34802.

Hesberg, C., Hänsch, R., Mendel, R. R., & Bittner, F. (2004). Tandem orientation of duplicated xanthine dehydrogenase genes from *Arabidopsis thaliana* differential gene expression and enzyme activities. *Journal of Biological Chemistry, 279*(14), 13547–13554.

Hille, R. (1996). The mononuclear molybdenum enzymes. *Chemical Reviews, 96*(7), 2757–2816.

Hille, R. (2002). Molybdenum and tungsten in biology. *Trends in Biochemical Sciences, 27*(7), 360–367.

Hilton J. C., & Rajagopalan K. V. (1996). Identification of the molybdenum cofactor of dimethyl sulfoxide reductase from *Rhodobacter sphaeroides* sp. *Denitrificans* as Bis(molybdopterin guanine dinucleotide)molybdenum. *Archives of Biochemestry and Biophysics*, *325*(1)139–143.

Hoff, T., Schnorr, K. M., Meyer, C., & Caboche, M. (1995). Isolation of two Arabidopsis cDNAs involved in early steps of molybdenum cofactor biosynthesis by functional complementation of *Escherichia coli* mutants. *Journal of Biological Chemistry, 270*(11), 6100–6107.

Ide, Y., Kusano, M., Oikawa, A., Fukushima, A., Tomatsu, H., Saito, K., et al. (2011). Effects of molybdenum deficiency and defects in molybdate transporter MOT1 on transcript accumulation and nitrogen/sulphur metabolism in *Arabidopsis thaliana*. *Journal of Experimental Botany*, *62*(4), 1483–1497.

Johnson, J. L. (2003). Prenatal diagnosis of molybdenum cofactor deficiency and isolated sulfite oxidase deficiency. *Prenatal Diagnosis, 23*(1), 6–8.

Johnson, J. L., & Rajagopalan, K. V. (1982). Structural and metabolic relationship between the molybdenum cofactor and urothione. *Proceedings of the National Academy of Sciences, 79*(22), 6856–6860.

Johnson, J. L., Bastian, N. R., & Rajagopalan, K. V. (1990). Molybdopterin guanine dinucleotide: a modified form of molybdopterin identified in the molybdenum cofactor of dimethyl sulfoxide reductase from *Rhodobacter sphaeroides* forma specialis denitrificans. *Proceedings of the National Academy of Sciences, 87*(8), 3190–3194.

Johnson, J. L., Rajagopalan, K. V., Lanman, J. T., Schutgens, R. B. H., Van Gennip, A. H., Sorensen, P., & Applegarth, D. A. (1991). Prenatal diagnosis of molybdenum cofactor deficiency by assay of sulphite oxidase activity in chorionic villus samples. *Journal of Inherited Metabolic Disease, 14*(6), 932–937.

Kloer, D. P., Hagel, C., Heider, J., & Schulz, G. E. (2006). Crystal structure of ethylbenzene dehydrogenase from *Aromatoleum aromaticum*. *Structure, 14*(9), 1377–1388.

Kuper, J., Llamas, A., Hecht, H. J., Mendel, R. R., & Schwarz, G. (2004). Structure of the molybdopterin-bound Cnx1G domain links molybdenum and copper metabolism. *Nature, 430*(7001), 803–806.

Leimkühler, S., Kern, M., Solomon, P. S., McEwan, A. G., Schwarz, G., Mendel, R. R., & Klipp, W. (1998). Xanthine dehydrogenase from the phototrophic purple bacterium *Rhodobacter capsulatus* is more similar to its eukaryotic counterparts than to prokaryotic molybdenum enzymes. *Molecular Microbiology, 27*(4), 853–869.

Leimkühler, S., Wuebbens, M. M., & Rajagopalan, K. V. (2001). Characterization of *Escherichia coli* MoeB and its involvement in the activation of molybdopterin synthase for the biosynthesis of the molybdenum cofactor. *Journal of Biological Chemistry, 276*(37), 34695–34701.

Lill, R., & Mühlenhoff, U. (2006). Iron-sulfur protein biogenesis in eukaryotes: Components and mechanisms. *Annual Review of Cell and Developmental Biology, 22*, 457–486.

Llamas, A., Mendel, R. R., & Schwarz, G. (2004). Synthesis of adenylated molybdopterin an essential step for molybdenum insertion. *Journal of Biological Chemistry, 279*(53), 55241–55246.

Magalon, A., Fedor, J. G., Walburger, A., & Weiner, J. H. (2011). Molybdenum enzymes in bacteria and their maturation. *Coordination Chemistry Reviews, 255*(9), 1159–1178.

Matthies, A., Nimtz, M., & Leimkühler, S. (2005). Molybdenum cofactor biosynthesis in humans: Identification of a persulfide group in the rhodanese-like domain of MOCS3 by mass spectrometry. *Biochemistry, 44*(21), 7912–7920.

Matthies, A., Rajagopalan, K. V., Mendel, R. R., & Leimkühler, S. (2004). Evidence for the physiological role of a rhodanese-like protein for the biosynthesis of the molybdenum cofactor in humans. *Proceedings of the National Academy of Sciences of the United States of America, 101*(16), 5946–5951.

McNicholas, P. M., Chiang, R. C., & Gunsalus, R. P. (1998). Anaerobic regulation of the *Escherichia coli* dmsABC operon requires the molybdate-responsive regulator ModE. *Molecular Microbiology, 27*(1), 197–208.

Mendel, R. R. (2013). The molybdenum cofactor. *Journal of Biological Chemistry, 288*(19), 13165–13172.

Mendel, R. R., & Bittner, F. (2006). Cell biology of molybdenum. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1763*(7), 621–635.

Mendel, R., & Schwarz, G. (2002). Biosynthesis and molecular biology of the molybdenum C of actor (Moco). *Metals ions in biological system: Volume 39: Molybdenum and tungsten: Their roles in biological processes*, 317.

Meneshian, A., & Bulkley, G. B. (2002). The physiology of endothelial xanthine oxidase: from urate catabolism to reperfusion injury to inflammatory signal transduction. *Microcirculation, 9*(3), 161–175.

Mercer, J. F. (2001). The molecular basis of copper-transport diseases. *Trends in Molecular Medicine, 7*(2), 64–69.

Millar, A. H., Liddell, A., & Leaver, C. J. (2001). Isolation and subfractionation of mitochondria from plants. *Methods in Cell Biology, 65*, 53–74.

Montalbini, P. (1992). Changes in xanthine oxidase activity in bean leaves induced by *Uromyces phaseoli* infection. *Journal of Phytopathology, 134*(1), 63–74.

Montalbini, P., & Della Torre, G. (1996). Evidence of a two-fold mechanism responsible for the inhibition by allopurinol of the hypersensitive response induced in tobacco by tobacco necrosis virus. *Physiological and Molecular Plant Pathology, 48*(4), 273–287.

Moorhead, G. B., Meek, S. E., Douglas, P., Bridges, D., Smith, C. S., Morrice, N., & MacKintosh, C. (2003). Purification of a plant nucleotide pyrophosphatase as a protein that interferes with nitrate reductase and glutamine synthetase assays. *European Journal of Biochemistry, 270*(6), 1356–1362.

Neumann, M., Mittelstädt, G., Iobbi-Nivol, C., Saggu, M., Lendzian, F., Hildebrandt, P., & Leimkühler, S. (2009). A periplasmic aldehyde oxidoreductase represents the first molybdopterin cytosine dinucleotide cofactor containing molybdo-flavoenzyme from *Escherichia coli*. *FEBS Journal, 276*(10), 2762–2774.

Partridge, J. D., Browning, D. F., Xu, M., Newnham, L. J., Scott, C., Roberts, R. E., et al. (2008). Characterization of the *Escherichia coli K-12* ydhYVWXUT operon: regulation by FNR, NarL and NarP. *Microbiology, 154*(2), 608–618.

Pastori, G. M., & del Rio, L. A. (1997). Natural senescence of pea leaves (an activated oxygen-mediated function for peroxisomes). *Plant Physiology, 113*(2), 411–418.

Pramanik, N. R., Ghosh, S., Raychaudhuri, T. K., Ray, S., Butcher, R. J., & Mandal, S. S. (2004). Synthesis, characterization and crystal structure of oxomolybdenum (VI) and (IV) complexes of some tridentate ONS donor ligands. *Polyhedron, 23*(9), 1595–1603.

Rajagopalan, K. V., & Johnson, J. L. (1992). The pterin molybdenum cofactors. *Journal of Biological Chemistry, 267*(15), 10199–10202.

Regulski, E. E., Moy, R. H., Weinberg, Z., Barrick, J. E., Yao, Z., Ruzzo, W. L., & Breaker, R. R. (2008). A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. *Molecular Microbiology, 68*(4), 918–932.

Reher, M., Gebhard, S., & Schönheit, P. (2007). Glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) and nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (GAPN), key enzymes of the respective modified Embden-Meyerhof pathways in the hyperthermophilic crenarchaeota *Pyrobaculum aerophilum* and *Aeropyrum pernix*. *FEMS Microbiology Letters, 273*(2), 196–205.

Reiss, J., & Johnson, J. L. (2003). Mutations in the molybdenum cofactor biosynthetic genes MOCS1, MOCS2, and GEPH. *Human Mutation, 21*(6), 569–576.

Reiss, J., Christensen, E., Kurlemann, G., Zabot, M. T., & Dorche, C. (1998). Genomic structure and mutational spectrum of the bicistronic MOCS1 gene defective in molybdenum cofactor deficiency type A. *Human Genetics, 103*(6), 639–644.

Rieder, C., Eisenreich, W., O'Brien, J., Richter, G., Götze, E., Boyle, P., et al. (1998). Rearrangement reactions in the biosynthesis of molybdopterin. *European Journal of Biochemistry*, *255*(1), 24–36.

Rivers, S. L., McNairn, E., Blasco, F., Giordano, G., & Boxer, D. H. (1993). Molecular genetic analysis of the moa operon of Escherichia coli K-12 required for molybdenum cofactor biosynthesis. *Molecular Microbiology, 8*(6), 1071–1081.

Rodríguez-Trelles, F., Tarrío, R., & Ayala, F. J. (2003). Convergent neofunctionalization by positive Darwinian selection after ancient recurrent duplications of the xanthine dehydrogenase gene. *Proceedings of the National Academy of Sciences, 100*(23), 13413–13417.

Romão, M. J. (2009). Molybdenum and tungsten enzymes: A crystallographic and mechanistic overview. *Dalton Transactions, 21*, 4053–4068.

Romão, M. J., Archer, M., Moura, I., Moura, J. J., LeGall, J., Engh, R., et al. (1995). Crystal structure of the xanthine oxidase-related aldehyde oxido-reductase from *D. gigas. Science, 270*(5239), 1170–1176.

Rothery, R. A., Workun, G. J., & Weiner, J. H. (2008). The prokaryotic complex iron–sulfur molybdoenzyme family. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1778*(9), 1897–1929.

Roy, R., & Adams, M. W. (2002). Characterization of a fourth tungsten-containing enzyme from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Journal of Bacteriology, 184*(24), 6952–6956.

Santamaria-Araujo, J. A., Fischer, B., Otte, T., Nimtz, M., Mendel, R. R., Wray, V., & Schwarz, G. (2004). The tetrahydropyranopterin structure of the sulfur-free and metal-free molybdenum cofactor precursor. *Journal of Biological Chemistry, 279*(16), 15994–15999.

Scheneider, F., Lowe, J., Huber, R., Schindelin, H., Kisker, C., & Knablein J. (1996). Crystal structure of dimethyl sulfoxide reductase from *Rhodobacter capsulatusat* 1.88 Å resolution. *Journal of Molecular Biology*, *263*(1), 53–69.

Schindelin, H., Kisker, C., Hilton, J., Rajagopalan, K. V., & Rees, D. C. (1996). Crystal structure of DMSO reductase: Redox-linked changes in molybdopterin coordination. *Science, 272*(5268), 1615–1621.

Schüttelkopf, A. W., Boxer, D. H., & Hunter, W. N. (2003). Crystal structure of activated ModE reveals conformational changes involving both oxyanion and DNA-binding domains. *Journal of Molecular Biology, 326*(3), 761–767.

Schwartz, S. H., Leon-Kloosterziel, K. M., Koornneef, M., & Zeevaart, J. A. (1997a). Biochemical characterization of the aba2 and aba3 mutants in *Arabidopsis thaliana*. *Plant Physiology, 114*(1), 161–166.

Schwartz, S. H., Tan, B. C., Gage, D. A., Zeevaart, J. A., & McCarty, D. R. (1997b). Specific oxidative cleavage of carotenoids by VP14 of maize. *Science, 276*(5320), 1872–1874.

Schwarz, G. (2005). Molybdenum cofactor biosynthesis and deficiency. *Cellular and Molecular Life Sciences CMLS, 62*(23), 2792–2810.

Schwarz, G., & Veldman, A. (2014). Molybdenum cofactor disorders. In Physician's guide to the diagnosis, treatment, and follow-up of inherited metabolic diseases (pp. 191–203). Berlin, Heidelberg: Springer.

Schwarz, G., Mendel, R. R., & Ribbe, M. W. (2009). Molybdenum cofactors, enzymes and pathways. *Nature, 460*(7257), 839–847.

Schwarz, G., Schulze, J., Bittner, F., Eilers, T., Kuper, J., Bollmann, G., et al. (2000). The molybdenum cofactor biosynthetic protein Cnx1 complements molybdate-repairable mutants, transfers molybdenum to the metal binding pterin, and is associated with the cytoskeleton. *The Plant Cell Online*, *12*(12), 2455–2471.

Self, W. T., Grunden, A. M., Hasona, A., & Shanmugam, K. T. (1999). Transcriptional regulation of molybdoenzyme synthesis in *Escherichia coli* in response to molybdenum: ModE-molybdate, a repressor of the *modABCD* (molybdate transport) operon is a secondary transcriptional activator for the *hyc* and *nar* operons. *Microbiology, 145*(1), 41–55.

Sigel, A., & Sigel, H. (Eds.). (2002). *Metals ions in biological system: Volume 39: Molybdenum and tungsten: Their roles in biological processes*. Boca Raton: CRC Press.

Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F., & Miller, N. E. (2001). Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Research, 29*(5), 1097–1106.

Solomon, P. S., Lane, I., Hanson, G. R., & Mcewan, A. G. (1997). Characterisation of the pterin molybdenum cofactor in dimethylsulfoxide reductase of *Rhodobacter capsulatus*. *European Journal of Biochemistry, 246*(1), 200–203.

Stallmeyer, B., Drugeon, G., Reiss, J., Haenni, A. L., & Mendel, R. R. (1999a). Human molybdopterin synthase gene: identification of a bicistronic transcript with overlapping reading frames. *The American Journal of Human Genetics, 64*(3), 698–705.

Stallmeyer, B., Schwarz, G., Schulze, J., Nerlich, A., Reiss, J., Kirsch, J., & Mendel, R. R. (1999b). The neurotransmitter receptor-anchoring protein gephyrin reconstitutes molybdenum cofactor biosynthesis in bacteria, plants, and mammalian cells. *Proceedings of the National Academy of Sciences, 96*(4), 1333–1338.

Teschner, J., Lachmann, N., Schulze, J., Geisler, M., Selbach, K., Santamaria-Araujo, J., et al. (2010). A novel role for *Arabidopsis* mitochondrial ABC transporter ATM3 in molybdenum cofactor biosynthesis. *The Plant Cell Online, 22*(2), 468–480.

Truglio, J. J., Theis, K., Leimkühler, S., Rappa, R., Rajagopalan, K. V., & Kisker, C. (2002). Crystal structures of the active and alloxanthine-inhibited forms of xanthine dehydrogenase from *Rhodobacter capsulatus*. *Structure, 10*(1), 115–125.

Unden, G., & Bongaerts, J. (1997). Alternative respiratory pathways of *Escherichia coli:* energetics and transcriptional regulation in response to electron acceptors. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, *1320*(3), 217–234.

Unden, G., Achebach, S., Holighaus, G., Tran, H. Q., Wackwitz, B., & Zeuner, Y. (2002). Control of FNR function of *Escherichia coli* by O_2 and reducing conditions. *Journal of Molecular Microbiology and Biotechnology, 4*(3), 263–268.

Wuebbens, M. M., & Rajagopalan, K. V. (1993). Structural characterization of a molybdopterin precursor. *Journal of Biological Chemistry, 268*(18), 13493–13498.

Wuebbens, M. M., & Rajagopalan, K. V. (1995). Investigation of the early steps of molybdopterin biosynthesis in *Escherichia coli* through the use of in vivo labeling studies. *Journal of Biological Chemistry, 270*(3), 1082–1087.

Yesbergenova, Z., Yang, G., Oron, E., Soffer, D., Fluhr, R., & Sagi, M. (2005). The plant Mo-hydroxylases aldehyde oxidase and xanthine dehydrogenase have distinct reactive oxygen species signatures and are induced by drought and abscisic acid. *The Plant Journal, 42*(6), 862–876.

Zhang, Y., & Gladyshev, V. N. (2008). Molybdoproteomes and evolution of molybdenum utilization. *Journal of Molecular Biology, 379*(4), 881–899.

Chapter 5
Nitrogen Fixation, a Molybdenum-Requiring Process

Abstract Nitrogen fixation is the most important input of biologically available nitrogen in Earth's ecosystems and is a metabolic ability possessed only by some Prokaryotes. To date four classes of nitrogenase enzymes have been characterized. Three nitrogenases are homologous enzymes with similar protein subunit composition and metal cofactor structure; these are the Mo-nitrogenase, V-nitrogenase, and Fe-only nitrogenase. How these three systems evolved and which of them first appeared on Earth is still under debate. The best studied system is the Fe-Mo-co based although several comparative analyses have been performed in past years.

In Chap. 4 only the class of pterin-based (Mo-co) molybdenum enzymes was discussed. However, there is an enzyme that uses a different form of Mo-cofactor (Fig. 5.1), the aforementioned nitrogenase, which catalyses the conversion of atmospheric nitrogen into ammonia, a process known as biological nitrogen fixation. Before entering nitrogenase's biochemical structure and the relative metabolic pathway, an overview of nitrogen fixation process will be provided.

5.1 Nitrogen Fixation in General

Nitrogen fixation is the most important input of biologically available nitrogen in Earth's ecosystems and is a metabolic ability possessed only by some Bacteria (Green Sulphur Bacteria, Firmicutes, Actinomycetes, Cyanobacteria and Proteobacteria) and Archaea, where it is mainly present in methanogens (Dixon and Kahn 2004). A recent analysis predicted that nearly 15 % of prokaryotic species with sequenced genomes are either known or potential diazotrophs (i.e. microbes able to fix nitrogen), a fraction much larger than commonly accepted (Dos Santos et al. 2012). Nitrogen fixation is a complex process with a high energetic cost and requiring the activity of several genes.

L. Presta et al., *Molybdenum Cofactors and Their role in the Evolution of Metabolic Pathways*, SpringerBriefs in Biometals, DOI 10.1007/978-94-017-9972-0_5

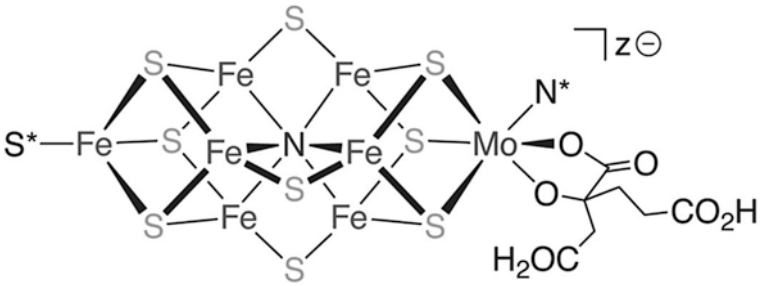

Fig. 5.1 Revised FeMo-cofactor structure derived from the most recent, high-resolution macromolecular structure determination. Reproduced from Lee and Holm (2003). With kind permission of © Proceedings of the National Academy of Sciences

To date four classes of nitrogenase enzymes have been characterized. Three nitrogenases are homologous enzymes with similar protein subunit composition and metal cofactor structure; these are the Mo-nitrogenase, V-nitrogenase, and Fe-only nitrogenase. The majority of present-day biological N_2 reduction is catalyzed by the Mo-dependent nitrogenase (encoded by *nif* genes), although V- and Fe-nitrogenase (encoded by *vnf* and *anf* genes, respectively) are important biological sources of fixed nitrogen in environments where Mo is present in limited amount. In fact the Mo-nitrogenase, which contains the iron-molybdenum cofactor or Fe-Mo-co, is the most commonly distributed nitrogenase; it is also the most efficient enzyme in the conversion of N_2 into NH_4^+. Alternative nitrogenases can be synthetized by diazotrophs, i.e. V-nitrogenase and/or Fe-only nitrogenase enzymes, which contain the FeV-co or Fe-co metalloclusters at their active sites, respectively. Regulation of alternative nitrogenases expression is dependent on the availability of molybdenum, vanadium or iron in the medium (Pau 2004). The three enzymes are oxygen-sensitive and they are irreparably damaged and deactivated by it. They show a high degree of conservation of structure, function and amino acid sequence across wide phylogenetic ranges. Interestingly, there are no reported diazotrophs lacking a Mo-nitrogenase and carrying uniquely an alternative nitrogenase. How these three systems evolved and which of them first appeared on Earth is still under debate. Several comparative analyses have been performed in past years although, to date, the best studied system is the Fe-Mo-co based.

The last type of nitrogenase is a Mo-nitrogenase phylogenetically unrelated to the other three classes and it has only been found in the bacterium *Streptomyces thermoautotrophicus*. This Mo-nitrogenase exhibits completely different biochemical features that consist of different protein composition, insensitivity to O_2, low Mg·ATP requirement, and a Mo-molybdopterin cytosine dinucleotide or Mo-MCD in the active site (Ribbe et al. 1997).

5.2 Biochemical Aspects

The overall reaction catalyzed by nitrogenase is usually depicted as follows:

$$\mathbf{N_2 + 8H^+ + 16MgATP + 8e^- \rightarrow 2NH_3 + H_2 + 16MgADP + 16Pi}\ \text{(inorganic phosphate)}.$$

This reaction not only represents the key entry point of reduced nitrogen into the global nitrogen cycle, but also embodies the formidable chemistry of breaking the triple bond of N_2 under ambient conditions (Hu and Ribbe 2011).

The FeMo-co nitrogenase (Fig. 5.2) is composed by two proteins, the iron (Fe) protein (encoded by *nifH*) and the molybdenum-iron (MoFe) protein (encoded by *nifD* and *nifK*). The homodimeric Fe protein is bridged by a single (4Fe–4S) cluster between the subunits and contains one ATP binding site within each subunit; whereas the $\alpha_2\beta_2$-tetrameric MoFe protein contains two unique clusters per αβ-dimer: the P-cluster (an (8Fe–7S) cluster), which is located at the α/β-subunit interface; and the Fe-Mo-co (a (Mo–7Fe–9S–X–homocitrate) cluster, where X is considered to be C, N or O), which is positioned within the α-subunit (Kim and Rees 1992). The catalysis of nitrogenase involves complex association/

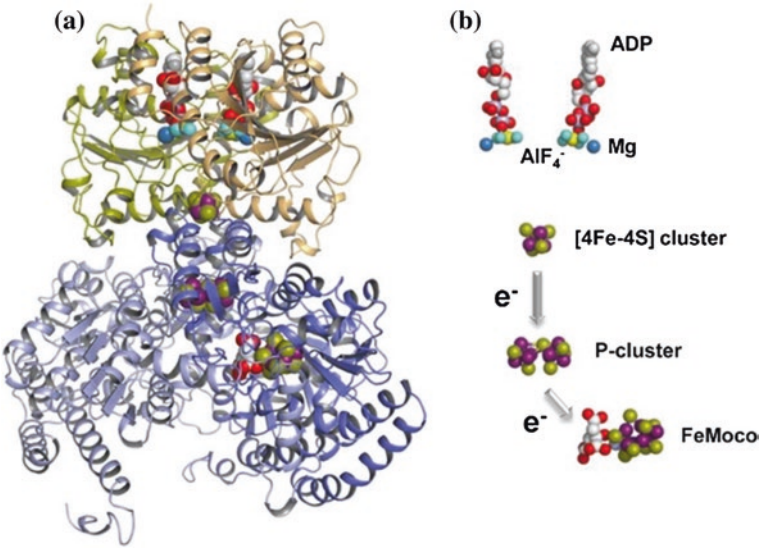

Fig. 5.2 X-ray crystal structure of half of the ADP·AlF_4^--stabilized Fe protein/MoFe protein complex (**a**) and the relative positions of components in the complex that are involved in the electron flow during catalysis (**b**). The identical subunits of Fe protein are shown in *yellow* and *orange*, and the α- and β-subunits of MoFe protein are shown in *light and dark blue*, respectively. The atoms of the components within the Fe protein/MoFe protein complex are colored as follows: Fe *purple*; S *yellow*; O *red*; C *gray*; Mg *dark blue*; Al *yellow*; F *light blue*. Reproduced from Hu and Ribbe (2011). With kind permission of © Coordination chemistry reviews

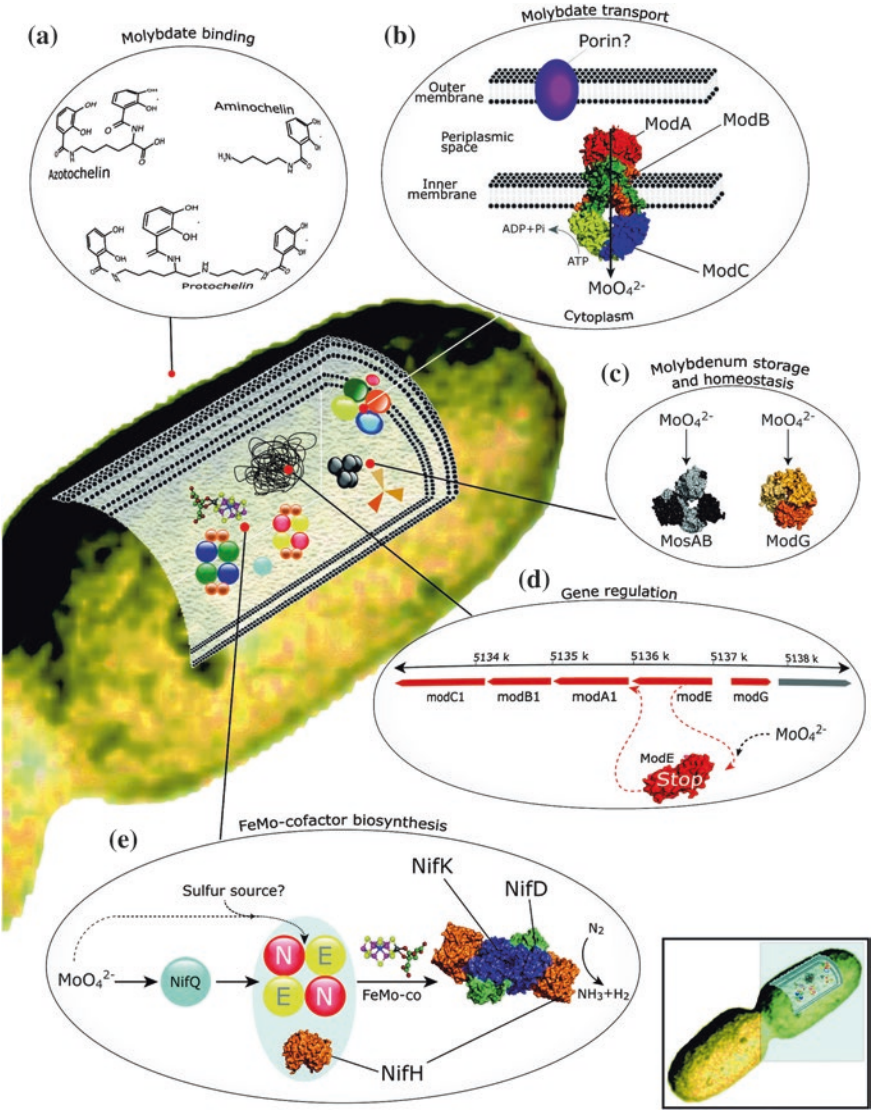

Fig. 5.3 Molybdenum trafficking for nitrogen fixation in the bacterium *Azotobacter vinelandii*. The figure shows a pathway of molybdenum towards the molybdenum-nitrogenase enzyme and the protein components involved in this pathway. Some of these proteins are exclusively dedicated to the nitrogenase biogenesis (e.g. NifQ and NifEN); some other have general roles in the metabolism of molybdenum (e.g. the ModABC molybdate transport system and the molybdenum-dependent transcriptional regulator ModE). Excreted siderophores with capacity to bind molybdate anions are also depicted. The *modABC* structure corresponds to the molybdate transport system of *Archaeoglobus fulgidus*. The ModE, NifH, and NifHDK structures shown are from *A. vinelandii*. Panel **e** shows a simplified FeMo-co biosynthetic pathway illustrating the two putative pathways for molybdenum incorporation into the Mo-nitrogenase cofactor. Reproduced from Hernandez et al. (2009). With kind permission of © Biochemistry

dissociation steps between the Fe protein and the MoFe protein, and the sequential, inter-protein transfer of electrons from the (4Fe–4S) cluster of the Fe protein, through the P-cluster, to the FeMoco of the MoFe protein, where substrate is reduced (Hu and Ribbe 2011).

The high level of complexity of nitrogenase's metallo clusters results in a (very) complex pathway for the assembly of nitrogenase and the insertion of the FeMo-co in the protein active site. The molybdenum trafficking for nitrogen fixation in the bacterium *A. vinelandii* is shown in Fig. 5.3.

In the free-living diazotroph *Klebsiella pneumoniae* at least 20 genes are involved in nitrogen fixation process (*nif* genes, Fig. 5.4) (Fani et al. 2000). In fact, apart from the catalytic components, additional gene's products are required to produce a fully functional enzyme. Indeed, several genes have been identified as

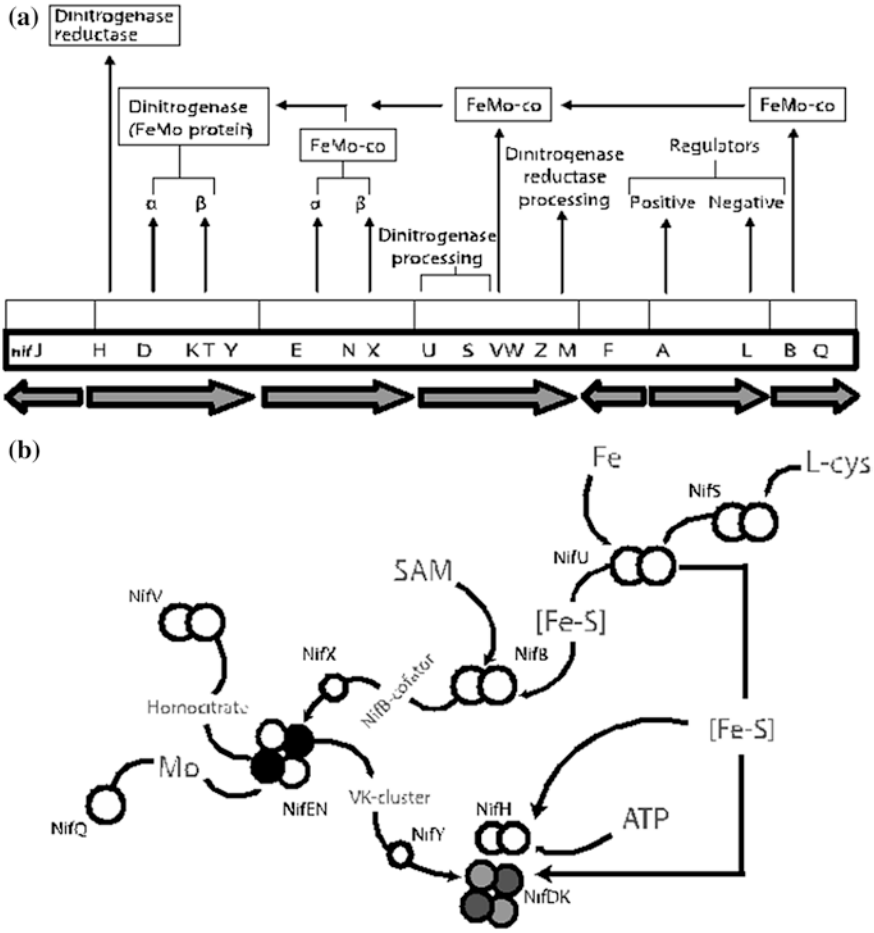

Fig. 5.4 **a** Organization of nif genes in *Klebsiella pneumonia* and **b** schematic representation of the nitrogen fixation process. Reproduced from Emiliani et al. (2010)

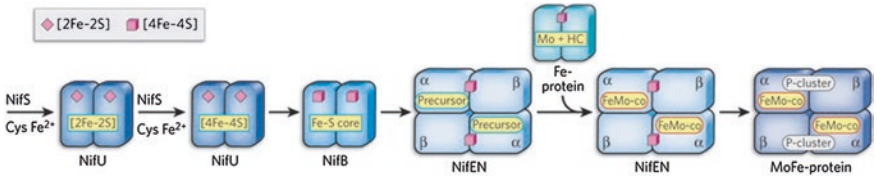

Fig. 5.5 Biosynthesis of FeMoco. NifS and NifU mobilize Fe and S for the sequential formation of $[Fe_2S_2]$ and $[Fe_4S_4]$ clusters, which are used as building blocks for the formation of a large Fe/S core on NifB. This Fe/S core is further processed into a molybdenum (Mo)- and homocitrate (HC)-free precursor, which can be converted to a mature FeMoco on NifEN upon Fe protein-mediated insertion of Mo and HC. Once the FeMoco is assembled on NifEN, it is delivered to its destined location in the MoFe protein. The permanent metal centers of the proteins are colored *gray* ($[Fe_2S_2]$ cluster *diamond*; $[Fe_4S_4]$ cluster *cube*; P-cluster *oval*), and the transient cluster intermediates are colored *yellow*. Reproduced from Schwarz et al. (2009). With kind permission of © Nature

being involved in this process (Fani et al. 2000 and references therein) although the number of proteins involved in the activation of nitrogenase seems to be species-specific and it varies according to the physiology of the organism and its environmental niche (Hamilton et al. 2011; Yan et al. 2010). In particular, FeMo-co biosynthesis requires, at least, the participation of *nifS*, *nifU*, *nifB*, *nifE*, *nifN*, *nifV*, *nifH*, *nifD* and *nifK* gene products. Based on genetic evidence, FeMo-co assembly is likely initiated by the actions of NifS and NifU (encoded by *nifS* and *nifU*) that generate small building blocks for FeMo-co assembly. NifS is a pyridoxal phosphate-dependent cysteine desulphurase, which forms a protein-bound cysteine persulphide that is subsequently donated to NifU for the sequential formation of small (2Fe–2S) and (4Fe–4S) clusters (Hu and Ribbe 2011). These small Fe–S fragments are then transferred to NifB (encoded by *nifB*) and processed into a FeMo-co core (with the help of NifV, Q and X). The last is then transferred to NifEN (encoded by *nifE* and *nifN*) and undergoes additional rearrangements before it is delivered to its target binding site in the MoFe protein through a protein-protein interaction (Hu and Ribbe 2011, Fig. 5.5).

5.3 How Did Nitrogen Fixation Evolve? Comparative Genomes Analyses Approach

The phylogenetic distribution of *nif* genes (i.e. the set of genes homologous to each of the 20 *nif* genes of another well-studied model organism, *Klebsiella pneumoniae*) in 842 completely sequenced prokaryotes (52 Archaea and 790 bacteria), revealed the presence of a common *core* of nitrogen fixation-related genes in a subset of 124 organisms (Emiliani et al. 2010). This *core* of genes is composed by only six *nif* genes (*nifHDKENB*), involved in nitrogenase, nitrogenase reductase and Fe–Mo Cofactor biosynthesis. Instead, all remaining *nif* genes display a

patchy phylogenetic distribution, thus revealing a complex evolutionary history of genes involved in the pathway. The finding of the same common *core* of genes has recently been confirmed by Dos Santos et al. (2012), who also proposed it as a novel in silico tool for the identification of additional diazotrophs in analyses performed in completely sequenced microbial genomes. Moreover, the finding of a common core of essential genes involved in the pathway might suggest that nitrogen fixation is an ancestral metabolic trait. Even if not proved, it is quite possible that only *nifHDKENB* genes were present in the genome of the LUCA (Last Universal Common Ancestor) community.

According to this idea, the *nifHDKENB* might represent a "universal core" for nitrogen fixation, whereas the other genes might have been differentially acquired during evolution in the different phylogenetic lineages. In fact, if nitrogen fixation had required other enzymes, their role might have been performed by others with low substrate specificity, in agreement with the Jensen hypothesis on the origin and evolution of metabolic pathways (Jensen 1976).

This idea is supported by analyses on the presence of *nif* genes homologs in the microbial domain revealing that most of the *nif* genes have in-paralogs (i.e. paralogs involved in the same pathway) and/or out-paralogs (i.e. paralogs involved in different pathways) as pointed out in Emiliani et al. (2010) (Fig. 5.6). The analysis did not retrieve any known paralogs for *nifW* (*nifO*), *nifT* (*fixU*), *nifQ* and *nifZ,* which are also missing from a large fraction of diazotrophs genomes. Eight *nif* genes (*nifAFHJLMSU*) are related, at a different extent, to proteins involved in other metabolic pathways (out-paralogs). NifS is related to a number of paralogs mainly involved in amino acid and carbon metabolisms. NifJ, a multidomain pyruvate:ferredoxin (flavodoxin) oxidoreductase, exhibited a large number of paralogs. Several proteins involved in Fe–Mo cofactor biosynthesis have paralogs in other cofactor biosyntheses.

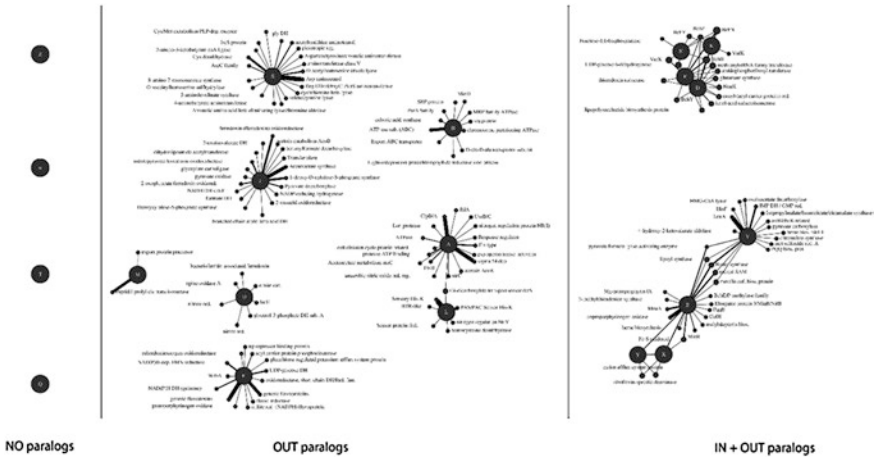

Fig. 5.6 In- and Out-paralogs network of nif genes. Nodes represent protein, links represent similarity values. Reproduced from Emiliani et al. (2010)

Eight Nif proteins share a significant degree of sequence similarity with proteins involved in other metabolic routes, and also with other *nif* genes products; this group can be further divided into two different clusters, the first of which includes *nifDKEN*, and the second being composed by *nifBXY* and *nifV*. NifB, NifX, and NifY share a common domain of about 90 aminoacids; moreover, *nifB* has an additional domain belonging to the S-adenosylmethionine (SAM) family, found in proteins that catalyze diverse reactions, including unusual methylations, isomerisation, sulphur insertion, ring formation, anaerobic oxidation and proteins' radical formation. Evidence exists that these proteins generate a radical species by reductive cleavage of SAM through an unusual Fe–S centre. The *nifV* and *nifB* genes are not directly linked although a sort of connection can be found in multidomain proteins sharing homology with NifV and NifB.

The evolutionary history of *nifDKEN* cluster has been investigate more in depth and more details are available to explain how it evolved (Fani et al. 2000).

Indeed, the two gene pairs *nifD-nifK* and *nifE-nifN* form a paralogous gene family, which code for nitrogenase and the tetrameric complex NifN2E2, respectively. Both *nifD-nifK* and *nifE-nifN* likely arose through duplications of an ancestral gene by a two-step model in which an ancestor gene underwent an in-tandem duplication event, giving rise to a bicistronic operon; this, in turn, duplicated, leading to the ancestors of the present-day *nifDK* and *nifEN* operons (Fani et al. 2000).

However, the role of these primordial enzymes is still under debate and it might have depended on the composition of the early atmosphere (Fig. 5.7). Accordingly,

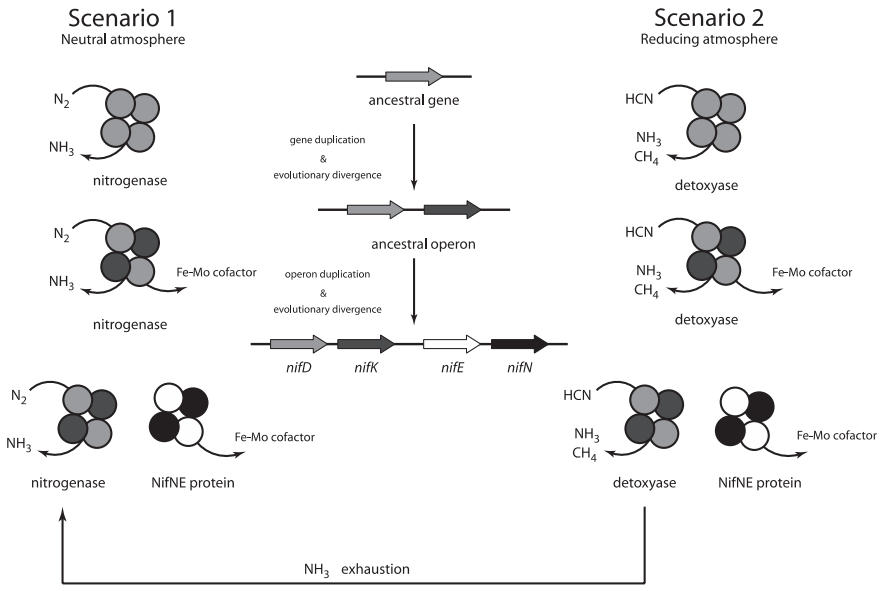

Fig. 5.7 Two possible scenarios depicted for the original function performed by the *nifDKEN* genes and their ancestor(s) gene(s). Reproduced from Fani et al. (2000). With kind permission of © Journal of Molecular Evolution

it can be accounted for by two alternative scenarios: (i) if we assume that O_2 was absent from the primordial atmosphere, an essential prerequisite for the appearance of (an ancestral) nitrogenase (since it is inactivated by free oxygen Fay 1992), then the appearance of nitrogenase would have represented a necessary event for the first cells, living in a planet whose atmosphere was neutral, containing dinitrogen, but not ammonia (first scenario). In fact, if the primitive microorganisms for their overall metabolism required ammonia, then its absence must have imposed a selective pressure favouring those cells that had evolved a system to synthesise ammonia from atmospheric dinitrogen.

Therefore, according to this scenario, the function of the ancestral enzyme might have been that of a slow, inefficient and with low substrate specificity "nitrogenase" thus being able to react with a wide range of compounds with a triple bond (Fani et al. 2000).

An alternative scenario would speculate an early reducing, free-ammonia rich atmosphere (Fig. 5.7). In those conditions, the evolution of a nitrogen fixation system would not have been a prerequisite because of the abundance of abiotically produced ammonia. Hence, why a nitrogenase in those days? The answer to this question relies in the catalytic properties of nitrogenase. In fact the enzyme is able to reduce also other molecules such as acetylene, hydrogen azide, hydrogen cyanide, or nitrous oxide, all of which contain a triple bond. Therefore, according to this second scenario (Fig. 5.7) the primitive enzyme encoded by the ancestor gene, would have been a detoxyase, an enzyme involved in detoxifying cyanides and other chemicals present in the primitive reducing atmosphere (Silver and Postgate 1973). This scenario implies that combined nitrogen progressive exhaustion would have imposed the refinement of the enzyme specificity. The last, very likely, should have modified and adapted to another triple-bond substrate, dinitrogen, and given to this change was selected and retained by some bacterial and archaeal lineages to enable survival in nitrogen-deficient environments. Finally, the decreasing of free ammonia and cyanides in the atmosphere triggered the evolution of the detoxyase toward nitrogenase, that might have been a common feature of all microbial life until photosynthesising cyanobacteria largely increased the oxygen concentration and burned cyanides.

Intriguingly, comparative genomes analyses also showed that genes coding for nitrogenase (*nifDK*) and nitrogenase reductase (*nifH*) are evolutionarily related to genes involved in bacteriochlorophyll biosynthesis (see below). Chlorophyll (Chl) and bacteriochlorophyll (Bchl) are the photochemically active reaction centre pigments for most of the extant photosynthetic organisms. During the synthesis of both Chl and Bchl, reduction of the tetrapyrrole ring system leads to the conversion of protochlorophyllide (Pchlide) into a chlorin. A second reduction, unique to the synthesis of Bchl, converts the chlorin into a bacteriochlorin. Two mechanisms for reducing the double bond in the fourth ring of protochlorophyllide are depicted. An enzymatic complex acts irrespective of the presence or absence of light and is thus termed "light-independent protochlorophyllide reductase". The second is a light-dependent reaction that utilizes the enzyme NADPH-protochlorophyllide oxidoreductase (Suzuki et al. 1997). In *Rhodobacter capsulatus*, the products of

three genes are required for each reduction: *bchL*, *bchN*, and *bchB* for the Pchlide reductase and *bchX*, *bchY*, and *bchZ* for the chlorin reductase (Burke et al. 1993b). Both enzymes are three-subunit complexes. Burke et al. (1993a, b) detected a significant degree of sequence similarity between BchlL, BchN, BchB, and BchX, BchY and BchZ, respectively, suggesting that the six genes represent two triads of paralogs and that the two enzymes are derived from a common three-subunit ancestral reductase. It was also found that the so-called "chlorophyll iron protein" subunits encoded by bchX, bchL, and chlL share a remarkable sequence similarity with the nitrogenase Fe proteins (Burke et al. 1993a). These findings led Burke et al. (1993b) to suggest that genes involved in bacteriochlorophyll biosynthesis and nitrogen fixation were related mechanistically, structurally and evolutionarily. Similarly to NifH protein, which serves as the unique electron donor for the nitrogenase complex, the products of *bchL* and *bchX* could serve as the unique electron donor into their respective catalytic subunits (BchB-BchN and BchY-BchZ). The idea of a common ancestry of *nifH*, *bchL* and *chlL* genes (Burke et al. 1993a; Fujita et al. 1993) has had an elegant experimental support by Cheng et al. (2005) who demonstrated in the photosynthetic eukaryote *Chlamydomonas reinhardtii* that NifH is able to partially complement the function of ChlL in the dark-dependent chlorophyll biosynthesis pathway.

Nitrogenases and carboxylases might have represented bacterial pre-adaptations, resulting in multigenic traits that were retained because leading to new selective advantages in altered environments. In fact, as depletion of abiotically produced organic matter occurred in early Earth environment, competition for the organic prerequisites for reproduction ensued. As the carboxylation and nitrogen-fixing functions were achieved, a new, abundant, and direct source of carbon and nitrogen for organic synthesis became available. The ability to take up atmospheric carbon and nitrogen would have been of great selective advantage (Margulis 1993). It is possible to propose a model (Fig. 5.8) for the origin and evolution of nitrogen fixation and bacterial photosynthesis based on multiple and successive paralogous duplications of an ancestral operon encoding an ancient reductase. The eight genes (*nifDKEN* and *bchYZNB*) are members of the same paralogous gene family, in that all of them are the descendants of a single ancestral gene. The model proposed posits the existence of an ancestral three-cistronic operon (Fig. 5.8) coding for an unspecific reductase. One might assume that this complex was (eventually) able to perform both carboxylation and nitrogen fixation. The following evolutionary steps might have been the duplication of the ancestral operon followed by an evolutionary divergence that led to the appearance of the ancestor of *nifH*, *nifDE*, and *nifKN* on one side, and *bchLX*, *bchNY* and *bchBZ* on the other one. In this way the two reductases narrowed their substrate specificity: one channelled toward nitrogen fixation and the other one toward photosynthesis. However, each of the two multicomplex proteins was able to perform at least two different reactions:

1. The ancestor of *nifDKEN*, was likely able to carry out the reduction of dinitrogen to ammonia and the synthesis of Fe–Mo cofactor (Fani et al. 2000).

2. The ancestor of protochlorophyllide- and chlorin-reductase performed both of the reactions that in the extant photosynthetic bacteria are carried out by two triads (BchLNB and BchXYZ, respectively).

The complete diversification of the function of the two heteromeric complexes was likely achieved thorough duplication of *nifDE* and *nifKN* ancestors and by the three-cistronic operon's duplication *bch*(LX)(NY)(BZ) followed by evolutionary divergence (Fig. 5.8). In our opinion, this idea may perfectly fit the Jensen's hypothesis.

Concerning the timing of the above reported evolutionary events (Fani et al. 2000) the two paralogous duplication events leading to *nifDK* and *nifEN* likely predated the appearance of the LUCA. Conversely, other authors (Raymond et al. 2004) have proposed a different scenario, according to which nitrogen fixation per se was invented by methanogenic Archaea and subsequently transferred, in at least three separate events, into bacterial lineages.

Differently from nitrogen fixation, tetrapyrrole-based photosynthesis occurs only in bacteria and bacterially derived chloroplasts, therefore it can be surmised that the appearance of photosynthesis should have not predated the divergence of Archaea and Bacteria.

Recently, a similar scenario has been proposed (Boyd et al. 2011). According to phylogenetic- and structure-based examinations of multiple nitrogenase proteins

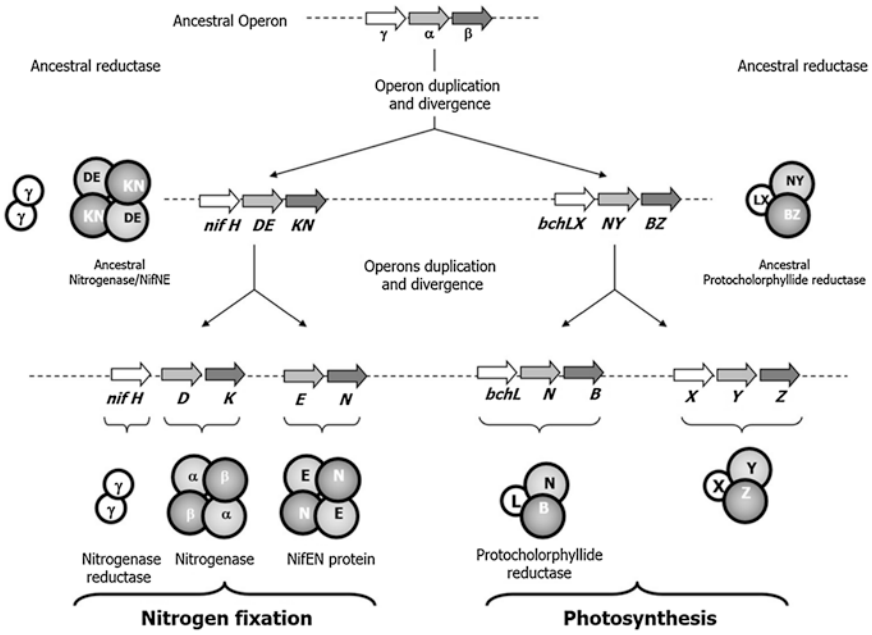

Fig. 5.8 Possible evolutionary model accounting for the evolutionary relationships between nif and bch genes. Reproduced from Emiliani et al. (2010)

these authors proposed a model for the emergence of nitrogenase whereby a gene encoding for an ancestral protein complex duplicated. The encoded protein should have possessed a cavity similar to that observed in the inferred structure of NflD (performing a nitrogenase-like reaction in Ni porphyrin F430 biosynthesis (Staples et al. 2007), the duplication led to the evolutionary precursor of BchN and NifD (including vanadium and Fe-only containing alternative nitrogenases, i.e. VnfD and AnfD). Serendipitously, metals (e.g., Fe) or metal clusters (e.g., 4Fe–4S) were bound in the cavity of the ancestor in a non-specific manner, resulting in an enzyme complex with altered reactivity, perhaps toward N_2 reduction (Boyd et al. 2011). In response to selective pressure of limited fixed nitrogen on early Earth, genes and associated gene's products were presumably recruited to improve the enzyme stepwise through the modification of the metal cofactor (Boyd et al. 2011).

5.4 Computational Studies on Mo-related Activities

In the last decade the availability of a large amount of data obtained from sequenced organisms has allowed the investigation of Mo utilization in prokaryotes and eukaryotes and its evolutionary changes. Comprehensive analyses of the occurrence and evolutionary trends of Mo-related traits can be carried out through bioinformatics tools by searching similar sequences of Mo uptake systems and Mo-dependent enzymes in the genomes of all so far sequenced organisms.

Recently several studies allowed the in silico identification of previously unknown features Mo-related such as the novel Mo-co-binding proteins that have been reported in both eukaryotes (mARC in pig mitochondria) (Havemeyer et al. 2006) and bacteria (YiiM and YcbX in *Escherichia coli*) (Kozmin et al. 2008).

These proteins share a significant degree of sequence homology with the C-terminal domain of eukaryotic Mo-co-sulfurase (MOSC) and show catalytic activity strictly dependent on Mo-co, suggesting that they may represent a previously unknown molybdoenzyme family. A study of 2011 by Yan Zhang et al. investigated the link between the MOSC motif and several known proteins highlighting interesting possible interaction between MO-systems and other pathways.

The MOSC domain is a superfamily of beta-strand-rich domains initially identified in the Mo-co sulfurase however it has subsequently been found in several other proteins from both prokaryotes and eukaryotes (Anantharaman and Aravind 2002). The MOSC domain of eukaryotic Mo-co sulfurase is involved in Mo-co binding with high affinity and its Mo-co carries a terminal sulfur ligand due to the catalytic activity of pyridoxal-5′-phosphate-dependent NifS-like domain (Wollers et al. 2008). On the other hand, Mo-co bound to the MOSC domain of mARC showed no terminal sulfur ligand. The function of the MOSC_N domain is unknown; however, it is predicted to adopt a beta barrel fold. Computational analysis allowed the distribution of these novel MOSC-containing molybdoproteins. It has been shown that all organisms containing MOSC-like proteins are Mo-co-utilizing organisms and that, in some organisms, genes for MOSC-like proteins are located close to

Mo-co biosynthesis components or molybdoenzymes such as MoaC and formate dehydrogenase. These findings indicate that MOSC-like proteins may be orthologs and they may serve as Mo-co chaperone involved in Mo-co transfer or storage. Further investigation may allow the complete understanding of interactions between Mo utilization and that of other trace elements, such as Fe and S.

The next chapter will focus on different methods and comparative genomics techniques through which computational studies can be carried out.

Bibliography

Anantharaman, V., & Aravind, L. (2002). MOSC domains: ancient, predicted sulfur-carrier domains, present in diverse metal–sulfur cluster biosynthesis proteins including Molybdenum cofactor sulfurases. *FEMS Microbiology Letters, 207*(1), 55–61.

Boyd, E. S., Hamilton, T. L., & Peters, J. W. (2011). An alternative path for the evolution of biological nitrogen fixation. *Frontiers in microbiology*, *2*.

Burke, D. H., Alberti, M., & Hearst, J. E. (1993a). The *Rhodobacter capsulatus* chlorin reductase-encoding locus, *bchA,* consists of three genes, *bchX, bchY,* and *bchZ. Journal of Bacteriology, 175*(8), 2407–2413.

Burke, D. H., Hearst, J. E., & Sidow, A. (1993b). Early evolution of photosynthesis: clues from nitrogenase and chlorophyll iron proteins. *Proceedings of the National Academy of Sciences, 90*(15), 7134–7138.

Cheng, Q., Day, A., Dowson-Day, M., Shen, G. F., & Dixon, R. (2005). *The Klebsiella pneumoniae* nitrogenase Fe protein gene (*nifH*) functionally substitutes for the *chlL* gene in *Chlamydomonas reinhardtii*. *Biochemical and biophysical research communications, 329*(3), 966–975.

Dixon, R., & Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nature Reviews Microbiology, 2*(8), 621–631.

Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C., & Dixon, R. (2012). Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics, 13*(1), 162.

Emiliani, G., Fondi, M., Liò, P., & Fani, R. (2010). *Evolution of metabolic pathways and evolution of genomes* (pp. 37–68). Netherlands: Springer.

Fani, R., Gallo, R., & Liò, P. (2000). Molecular evolution of nitrogen fixation: the evolutionary history of the *nifD*, *nifK*, *nifE*, and *nifN* genes. *Journal of Molecular Evolution, 51*(1), 1–11.

Fay, P. (1992). Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiological Reviews, 56*(2), 340.

Fujita, Y., Matsumoto, H., Takahashi, Y., & Matsubara, H. (1993). Identification of a *nifDK*-like gene (ORF467) involved in the biosynthesis of chlorophyll in the cyanobacterium *Plectonema boryanum*. *Plant and Cell Physiology, 34*(2), 305–314.

Hamilton, T. L., Ludwig, M., Dixon, R., Boyd, E. S., Dos Santos, P. C., Setubal, J. C., et al. (2011). Transcriptional profiling of nitrogen fixation in *Azotobacter vinelandii*. *Journal of bacteriology*, *193*(17), 4477–4486.

Havemeyer, A., Bittner, F., Wollers, S., Mendel, R., Kunze, T., & Clement, B. (2006). Identification of the missing component in the mitochondrial benzamidoxime prodrug-converting system as a novel molybdenum enzyme. *Journal of Biological Chemistry, 281*(46), 34796–34802.

Hernandez, J. A., George, S. J., & Rubio, L. M. (2009). Molybdenum trafficking for nitrogen fixation. *Biochemistry, 48*(41), 9711–9721.

Hu, Y., & Ribbe, M. W. (2011). Biosynthesis of nitrogenase FeMoco. *Coordination Chemistry Reviews, 255*(9), 1218–1224.

Jensen, R. A. (1976). Enzyme recruitment in evolution of new function. *Annual Reviews in Microbiology, 30*(1), 409–425.

Kim, J., & Rees, D. C. (1992). Structural models for the metal centers in the nitrogenase molybdenum-iron protein. *Science, 257*(5077), 1677–1682.

Kozmin, S. G., Leroy, P., Pavlov, Y. I., & Schaaper, R. M. (2008). YcbX and yiiM, two novel determinants for resistance of *Escherichia coli* to N-hydroxylated base analogues. *Molecular Microbiology, 68*(1), 51–65.

Lee, S. C., & Holm, R. H. (2003). Speculative synthetic chemistry and the nitrogenase problem. *Proceedings of the National Academy of Sciences, 100*(7), 3595–3600.

Margulis, L. (1993). Symbiosis in cell evolution: Microbial communities in the Archean and Proterozoic eons.

Pau, R. N. (2004). Molybdenum uptake and homeostasis. In W. Klipp, B. Masepohl, J. R. Gallon, & W. E. Newton (Eds.), *Genetics and regulation of nitrogen fixation in free-living bacteria* (pp. 225–256). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Raymond, J., Siefert, J. L., Staples, C. R., & Blankenship, R. E. (2004). The natural history of nitrogen fixation. *Molecular Biology and Evolution, 21*(3), 541–554.

Ribbe, M., Gadkari, D., & Meyer, O. (1997). N2 Fixation by *Streptomyces thermoautotrophicus* involves a molybdenum-dinitrogenase and a manganese-superoxide oxidoreductase that couple N_2 reduction to the oxidation of superoxide produced from O_2 by a molybdenum-CO dehydrogenase. *Journal of Biological Chemistry, 272*(42), 26627–26633.

Schwarz, G., Mendel, R. R., & Ribbe, M. W. (2009). Molybdenum cofactors, enzymes and pathways. *Nature, 460*(7257), 839–847.

Silver, V. S., & Postgate, J. R. (1973). Evolution of asymbiotic nitrogen fixation. *Journal of Theoretical Biology, 56*, 340–373.

Staples, C. R., Lahiri, S., Raymond, J., Von Herbulis, L., Mukhophadhyay, B., & Blankenship, R. E. (2007). Expression and association of group IV nitrogenase NifD and NifH homologs in the non-nitrogen-fixing archaeon *Methanocaldococcus jannaschii*. *Journal of Bacteriology, 189*(20), 7392–7398.

Suzuki, J. Y., Bollivar, D. W., & Bauer, C. E. (1997). Genetic analysis of chlorophyll biosynthesis. *Annual Review of Genetics, 31*(1), 61–89.

Wollers, S., Heidenreich, T., Zarepour, M., Zachmann, D., Kraft, C., Zhao, Y., & Bittner, F. (2008). Binding of sulfurated molybdenum cofactor to the C-terminal domain of ABA3 from *Arabidopsis thaliana* provides insight into the mechanism of molybdenum cofactor sulfuration. *Journal of Biological Chemistry, 283*(15), 9642–9650.

Yan, Y., Ping, S., Peng, J., Han, Y., Li, L., Yang, J., & Lin, M. (2010). Global transcriptional analysis of nitrogen fixation and ammonium repression in root-associated *Pseudomonas stutzeri A1501*. *BMC Genomics, 11*(1), 11.

Chapter 6
Protocols and Methods for the in Silico Reconstruction of the Origin and Evolution of Metabolic Pathways

Abstract It is possible to reconstruct the origin and evolution of metabolic pathways by inferring useful hints from the analyses of metabolic pathways existing in contemporary cells. Several bioinformatic tools allow the comparison of gene and genomes from organisms belonging to the three cell domains, giving rise to comparative genomics. Moreover the storage, the interpretation and the integration of several sources of information (gene structure and organization, gene regulation, protein-protein interactions) make possible to infer phylogenetic relationships among organisms, leading to a new approach referred to as phylogenomics. A short description of several tools is reported.

How can the origin ad evolution of metabolic pathways be studied and reconstructed? By assuming that useful hints may be inferred from the analysis of metabolic pathways existing in contemporary cells, important insights on the evolutionary development of microbial metabolic pathways can be obtained by the use of bioinformatic tools, allowing the comparison of gene and genomes from organisms belonging to the three cell domains (Archaea, Bacteria and Eukarya). This approach takes advantage of the availability of the phylogenetic relationships among (micro)organisms, and possibly on the existence of different structure and organization exhibited by orthologous genes. Beside, the more ancient is a pathway, the more information can be retrieved from this comparative analysis. Such kind of approach has become feasible since, in recent years, saw a dramatic increase in genomics data deriving from organisms belonging to all of the three known domains of life. By the way, the use of bioinformatic tools allowed the storage and the interpretation of several sources of information (gene structure and organization, gene regulation, protein-protein interactions) and, probably more importantly, their integration, a fundamental step for the global understanding of genomes properties and dynamics. This approach is usually referred to as comparative genomics. Combining data gained from comparative genomics with evolutionary studies of different species (i.e. phylogenetic inference), results in a new kind of approach, referred to as phylogenomics. A typical phylogenomics pipeline is illustrated in Fig. 6.1.

L. Presta et al., *Molybdenum Cofactors and Their role in the Evolution of Metabolic Pathways*, SpringerBriefs in Biometals, DOI 10.1007/978-94-017-9972-0_6

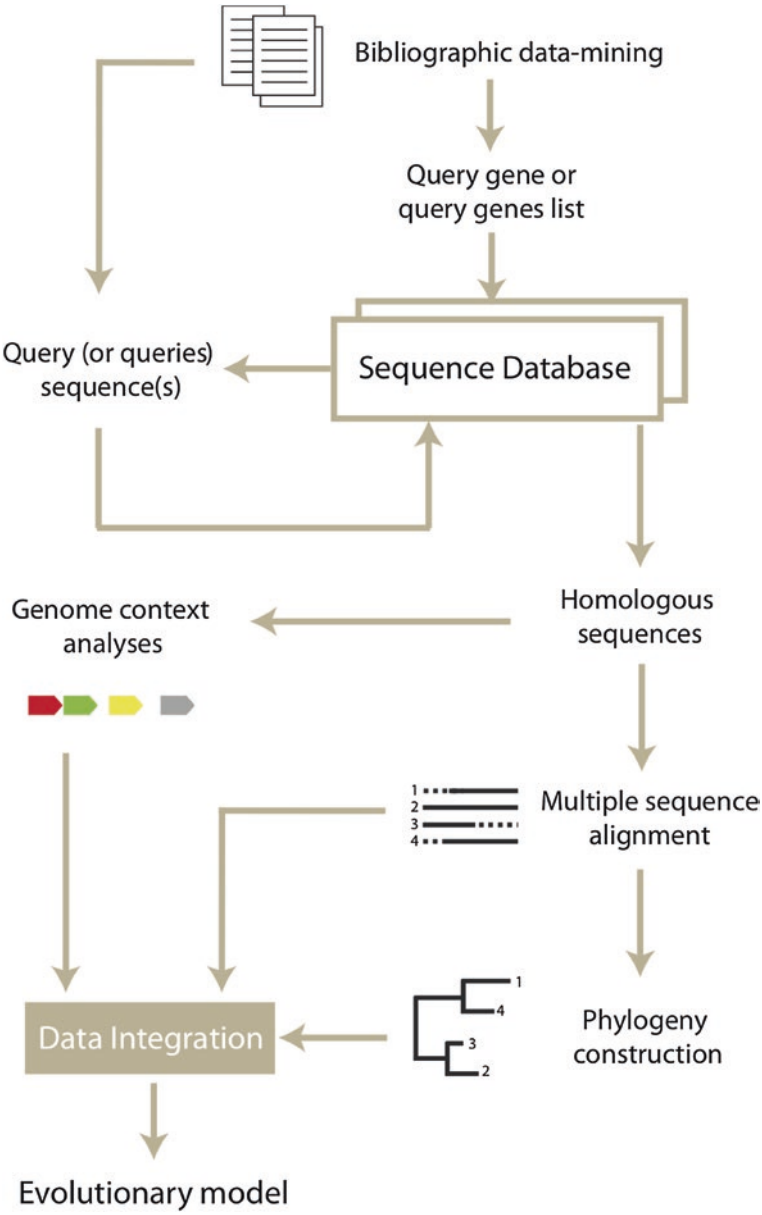

Fig. 6.1 A schematic representation of a typical phylogenetic pipeline

This way of investigating the evolutionary history of genes introduced several advantages; in fact, adopting a genome-scale approach theoretically overcomes incongruence derived from molecular phylogenies based on single genes mainly because (i) non-orthologous comparison (i.e. the comparison of those genes

erroneously defined as orthologous) is much more misleading when the analysis is performed on a single gene, whereas it is probably buffered in a multigene analysis and (ii) stochastic error naturally vanishes when more and more genes are considered. At the same time it is useful to recur to bioinformatics methods in order to detect new genes or proteins or to predict their function.

6.1 Browsing Microbial Genomes

At present, hundreds of microbial genomes have been sequenced, and hundreds more are currently in the pipeline. Furthermore, functional genomic studies have generated a large and growing body of experimental results for many different organisms belonging to the known domains of life. However, this whole body of data would reveal almost useless if not stored in a proper manner. To this purpose a growing number of public databases have been developed in recent years, usually providing also user-friendly tools for their interrogation. These tools, despite not allowing automatized large-scale phylogenomic analyses, often represent their first preliminary (and useful) step. This is the case for example of MicrobesOnLine [http://www.microbesonline.org (Alm et al. 2005; Dehal et al. 2010)], which embeds both structural and functional data on a large (almost 3000) dataset of completely sequenced genomes. These data are retrieved from a wide range of other specific databases (including KEGG, GeneOntology, RefSeq).

Interestingly, MicrobesOnLine also allows to interactively explore the neighborhood of any given gene, hence allowing, for example, a first analysis of the gene organization of a given metabolic pathway. Similarly, IMG (Chu et al. 2013) provides users some linked tools to support comparative microbial genes, genomes, and metagenomics analysis, including COG, KEGG, Pfam, InterPro, and the Gene Ontology. Consequently, thanks also to the graphical user interface IMG is particularly suited for nonexperienced bioinformaticians which want to perform comparative genome analyses (Checcucci and Mengoni 2015).

The same task can be pursued adopting also operonDB web service (http://odb.kuicr.kyoto-u.ac.jp/), (Pertea et al. 2009) aiming at collecting all known operons (derived from the literature and from publicly available database) in multiple species and to offer a system to predict operons by user definitions. Several other web sites and software tools have been described that assist in the annotation and exploration of comparative genomic data. The Prolinks (Bowers et al. 2004) and STRING (Jensen et al. 2009) databases offer convenient tools for browsing predicted functional associations among proteins. String, in particular imports protein association knowledge not only from databases of physical interactions, but also from databases of curated biological pathway knowledge. A number resources are included in the current release [MINT (Ceol et al. 2009), HPRD (Prasad et al. 2009), DIP (Xenarios et al. 2002), BioGRID (Stark et al. 2008), KEGG (Kanehisa and Goto 2000) and Reactome (Matthews et al. 2009) IntAct (Hermjakob et al. 2004), EcoCyc (Keseler et al. 2009)]. Furthermore, this set of previously known

and well-described interactions is then complemented by interactions that are predicted computationally, specifically for STRING, using a number of prediction algorithms (Jensen et al. 2009).

6.2 Orthologs Identification

Genomics data is a fundamental step for addressing the topic of the evolution of metabolic pathways, and strictly depends on a correct identification of orthologous proteins shared by different genomes. This field has been greatly developed in recent years and, paradoxically, the extant challenge seems not to be the lack of orthology predictions, but the right choice within the plethora of methods and databases that have been recently implemented (Gabaldon et al. 2009). The identification of orthologs between two genomes often relies on the so-called bidirectional best-hit (BBH) criterion, a reiteration of the BLAST algorithm (Altschul et al. 1997): two proteins, a and b, from genomes A and B respectively, are orthologs if a is the best-hit (i.e. the most similar) of b in genome A and vice versa. For three or more genomes, groups of orthologous sequences can be constructed by extending the BBH relationships with a clustering algorithm. This approach has led to the assembly of pre-compiled databases embedding groups of orthologous proteins, such as COG or KEGG-related systems (KOBAS and KAAS). Moreover, several other algorithms have been developed to fulfill this tasks, including Ncut (Abascal and Valencia 2002), Rio, (Zmasek and Eddy 2002), Outgroup Conditioned Score (OCS) (Cotter et al. 2002) or OrthoParaMap (Cannon and Young 2003). Recent advancements showed that clustering techniques applied to matrices storing pair-wise similarities perform quite well (Brilli et al. 2008). These algorithms work either on the grouping of weakly similar homologs or on the identification of protein domains. The most widespread are: (i) orthoMCL (Li et al. 2003), which adopts a Markov Clustering algorithm (previously implemented in tribeMCL (Enright et al. 2002), (ii) Ortholuge (Fulton et al. 2006) that aims at identifying orthologs by comparing proteins and species phylogenetic trees and, lastly, (iii) InParanoid (O'Brien et al. 2005) that relies on a similar flowchart. All these orthologs' identification methods have been recently tested on a dataset of proteins from different species previously characterized using functional genomics data, such as expression data and protein interaction data (Hulsen et al. 2006). Results have shown that InParanoid software seems the best orthologs' identification method in terms of identifying functionally equivalent proteins in different species (Hulsen et al. 2006).

6.3 Multiple Sequence Alignments

In a phylogenetic analysis workflow (but also when interested, for example, in structure modelling, functional site prediction and sequence database searching) a key step (usually following the correct orthologs retrieval procedure) consists in comparing those residues with inferred common evolutionary origin or structural/functional

equivalence in the whole sequence dataset. This task is fulfilled through multiple sequence alignment (MSA), that is arranging homolog protein sequences into a rectangular array with the goal that residues in a given column are homologous (derived a single position in an ancestral sequence), superposable (in a rigid local structural alignment) or play a common functional role. Although these three criteria are essentially equivalent for closely related proteins, sequence, structure and function diverge over evolutionary time and different criteria may result in different alignments (Edgar and Batzoglou 2006). Many approximate algorithms have been developed for multiple sequence alignments, including the commonly used progressive alignment technique (Pei 2008). This greedy heuristic assembly algorithm involves estimating a guide tree (rooted binary tree) from unaligned sequences and then incorporating the sequences into the MSA with a pairwise alignment algorithm while following the tree topology. The scoring schemes used by the pairwise alignment algorithm are arguably the most influential component of the progressive algorithm. They can be divided in two categories, that is matrix- and consistency-based algorithms. Matrix-based algorithms such as ClustalW (Thompson et al. 2002), MUSCLE (Edgar 2004), and Kalign (Lassmann and Sonnhammer 2005) use a substitution matrix to assess the cost of matching two symbols or two profiled columns (Notredame 2007). Conversely, consistency-based schemes incorporate a larger share of information into the evaluation. This result is achieved by using an approach initially developed for T-Coffee (Notredame et al. 2000) and inspired by Dialign overlapping weights (Morgenstern et al. 1998; Subramanian et al. 2005). Its principle is to compile a collection of pairwise global and local alignments (primary library) and to use this collection as a position-specific substitution matrix during a regular progressive alignment. The aim is to deliver a final MSA as consistent as possible with the alignments contained in the library. Many extant algorithms are based on this approach such as PCMA (Pei et al. 2003), ProbCons (adopting a Bayesian framework) (Do et al. 2005), MUMMALS (Pei and Grishin 2007). Sequence and structural databases are expanding rapidly owing to genome sequencing projects and structural genomics initiatives, offering helpful sources to further improve multiple protein sequence alignments. Structural additional information, for example known 3-dimensional (3D) structures, can be exploited in some multiple alignment methods. In fact, since structures are generally more conserved than sequences, structural information is also valuable for aligning sequences. Several MS algorithm have started implementing this source of information, and they include 3DCoffee (Poirot et al. 2004) and FUGUE (Shi et al. 2001). Recently, the Expresso server (Armougom et al. 2006) extended the 3DCoffee method by automatically identifying highly similar 3D structural templates for target sequences and using structural alignments for consistency-based alignments.

6.4 Phylogeny Reconstruction

Understanding microbial evolution is essential for gathering information on the most ancient events in the history of Life on our planet (Gribaldo and Brochier 2009) as well as on the extant relationships between whole microbial communities. This task

implies the use of molecular phylogeny techniques that is the study of phylogenies and processes of evolution by the analysis of DNA or amino acid sequence data (Whelan and Goldman 2001). Although parsimony and distance-based methods are widely used, the most statistically robust approach is to consider the problem in a likelihood framework and use accurate models of evolution (Brilli et al. 2008). It is known (Whelan and Goldman 2001), in fact, that disadvantages of distance methods include the inevitable loss of evolutionary information when a sequence alignment is converted to pairwise distances, and the inability to deal with models containing parameters for which the values are not known a priori. Concerning maximum parsimony (MP), this approach selects and outputs the tree (or trees) that require the fewest evolutionary changes and is reasonably confident when the number of changes per sequence position is relatively small (Steel and Penny 2000). However, as more-divergent sequences are to be analyzed, the degree of homoplasy (i.e. parallel, convergent, reversed or superimposed changes) increases and MP tree reconstruction might be misleading since this method has no adequate means to deal with this (Whelan and Goldman 2001). Conversely, Maximum likelihood (ML) approaches take the hypothesis (the tree topology) that maximizes the likelihood of the data (the sequence alignment) in the light of an evolutionary model. A great attraction of this approach is the ability to perform robust statistical hypothesis tests and to use modern statistical techniques such as hidden Markov models, Markov chain Monte Carlo and Bayesian inference (Ewens and Grant 2005; Shoemaker et al. 1999). The ML framework also allows each site of the alignment to evolve with different replacement patterns, and with different substitution rates in all branches of the tree (Whelan and Goldman 2001) as in real proteins, where slowly evolving sites are generally functionally or structurally constrained, while variable sites are likely to be less important for protein function. The ML approach (including its variants as the Bayesian framework) has been included in a number of different packages, such as Phylip (http://evolution.gs.washington.edu/phylip.html) PAUP* (http://paup.csit.fsu.edu/) MEGA http://www.megasoftware.net/mega.html, (Tamura et al. 2008), PAML [http://abacus.gene.ucl.ac.uk/software/paml.html, (Yang 1997)], mrBayes (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003) and phyML (Guindon and Gascuel 2003).

Bibliography

Abascal, F., & Valencia, A. (2002). Clustering of proximal sequence space for the identification of protein families. *Bioinformatics, 18*(7), 908–921.

Alm, E. J., Huang, K. H., Price, M. N., Koche, R. P., Keller, K., Dubchak, I. L., & Arkin, A. P. (2005). The MicrobesOnline web site for comparative genomics. *Genome Research, 15*(7), 1015–1022.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research, 25*(17), 3389–3402.

Armougom, F., Moretti, S., Keduas, V., & Notredame, C. (2006). The iRMSD: A local measure of sequence alignment accuracy using structural information. *Bioinformatics, 22*(14), e35–e39.

Bowers, P. M., Pellegrini, M., Thompson, M. J., Fierro, J., Yeates, T. O., & Eisenberg, D. (2004). Prolinks: A database of protein functional linkages derived from coevolution. *Genome Biology, 5*(5), R35.

Brilli, M., Mengoni, A., Fondi, M., Bazzicalupo, M., Liò, P., & Fani, R. (2008). Analysis of plasmid genes by phylogenetic profiling and visualization of homology relationships using Blast2Network. *BMC Bioinformatics, 9*(1), 551.

Cannon, S. B., & Young, N. D. (2003). OrthoParaMap: Distinguishing orthologs from paralogs by integrating comparative genome data and gene phylogenies. *BMC Bioinformatics, 4*(1), 35.

Ceol, A., Aryamontri, A. C., Licata, L., Peluso, D., Briganti, L., Perfetto, L., & Cesareni, G. (2009). MINT, the molecular interaction database: 2009 update. *Nucleic Acids Research*, gkp98.

Checcucci, A., & Mengoni, A. (2015). The integrated microbial genome resource of analysis in bacterial pangenomics (pp. 289–295). Springer, New York.

Chu, K., Markowitz, V. M., Chen, I. M. A., Palaniappan, K., Szeto, E., Pillay, M., & Kyrpides, N. C. (2013). IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Research*, gkt963.

Cotter, P. J., Caffrey, D. R., & Shields, D. C. (2002). Improved database searches for orthologous sequences by conditioning on outgroup sequences. *Bioinformatics, 18*(1), 83–91.

Dehal, P. S., Joachimiak, M. P., Price, M. N., Bates, J. T., Baumohl, J. K., Chivian, D., et al. (2010). MicrobesOnline: An integrated portal for comparative and functional genomics. *Nucleic Acids Research, 38*(suppl 1), D396–D400.

Do, C. B., Mahabhashyam, M. S., Brudno, M., & Batzoglou, S. (2005). ProbCons: Probabilistic consistency-based multiple sequence alignment. *Genome Research, 15*(2), 330–340.

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research, 32*(5), 1792–1797.

Edgar, R. C., & Batzoglou, S. (2006). Multiple sequence alignment. *Current Opinion in Structural Biology, 16*(3), 368–373.

Enright, A. J., Van Dongen, S., & Ouzounis, C. A. (2002). An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research, 30*(7), 1575–1584.

Ewens, W. J., & Grant, G. R. (2005). *Statistical methods in bioinformatics: An introduction* (Vol. 746867830). New York: Springer.

Fulton, D. L., Li, Y. Y., Laird, M. R., Horsman, B. G., Roche, F. M., & Brinkman, F. S. (2006). Improving the specificity of high-throughput ortholog prediction. *BMC Bioinformatics, 7*(1), 270.

Gabaldón, T., Silla-Martínez, J. M., & Capella-Gutiérrez, S. (2009). TrimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics, 25*(15), 1972–1973.

Gribaldo, S., & Brochier, C. (2009). Phylogeny of prokaryotes: Does it exist and why should we care? *Research in Microbiology, 160*(7), 513–521.

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology, 52*(5), 696–704.

Hermjakob, H., Montecchi-Palazzi, L., Bader, G., Wojcik, J., Salwinski, L., Ceol, A., & Apweiler, R. (2004). The HUPO PSI's molecular interaction format—A community standard for the representation of protein interaction data. *Nature Biotechnology, 22*(2), 177–183.

Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics, 17*(8), 754–755.

Hulsen, T., Huynen, M. A., de Vlieg, J., & Groenen, P. M. (2006). Benchmarking ortholog identification methods using functional genomics data. *Genome Biology, 7*(4), R31.

Jensen, L. J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., et al. (2009). STRING 8—A global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Research, 37*(suppl 1), D412–D416.

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research, 28*(1), 27–30.

Keseler, I. M., Bonavides-Martínez, C., Collado-Vides, J., Gama-Castro, S., Gunsalus, R. P., Johnson, D. A., & Karp, P. D. (2009). EcoCyc: A comprehensive view of Escherichia coli biology. *Nucleic Acids Research, 37*(suppl 1), D464–D470.

Lassmann, T., & Sonnhammer, E. L. (2005). Kalign—An accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics, 6*(1), 298.

Li, L., Stoeckert, C. J., & Roos, D. S. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research, 13*(9), 2178–2189.

Matthews, L., Gopinath, G., Gillespie, M., Caudy, M., Croft, D., de Bono, B., et al. (2009). Reactome knowledgebase of human biological pathways and processes. *Nucleic Acids Research, 37*(suppl 1), D619–D622.

Morgenstern, B., Frech, K., Dress, A., & Werner, T. (1998). DIALIGN: Finding local similarities by multiple sequence alignment. *Bioinformatics, 14*(3), 290–294.

Notredame, C. (2002). Recent progress in multiple sequence alignment: A survey. *Pharmacogenomics, 3*(1), 131–144.

Notredame, C. (2007). Recent evolutions of multiple sequence alignment algorithms. *PLoS Computational Biology, 3*(8), e123.

Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, *302*(1), 205–217.

O'Brien, K. P., Remm, M., & Sonnhammer, E. L. (2005). Inparanoid: A comprehensive database of eukaryotic orthologs. *Nucleic Acids Research, 33*(suppl 1), D476–D480.

Pei, J. (2008). Multiple protein sequence alignment. *Current Opinion in Structural Biology, 18*(3), 382–386.

Pei, J., & Grishin, N. V. (2007). PROMALS: Towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics, 23*(7), 802–808.

Pei, J., Sadreyev, R., & Grishin, N. V. (2003). PCMA: Fast and accurate multiple sequence alignment based on profile consistency. *Bioinformatics, 19*(3), 427–428.

Pertea, G., Zimin, A. V., Delcher, A. L., Florea, L., Kelley, D. R., Schatz, M. C., et al. (2009). A whole-genome assembly of the domestic cow, Bos taurus. *Genome Biology, 10*(4), R42.

Poirot, O., Suhre, K., Abergel, C., O'Toole, E., & Notredame, C. (2004). 3DCoffee@ igs: A web server for combining sequences and structures into a multiple sequence alignment. *Nucleic Acids Research, 32*(suppl 2), W37–W40.

Prasad, T. K., Kandasamy, K., & Pandey, A. (2009). Human protein reference database and human proteinpedia as discovery tools for systems biology. In *Reverse Chemical Genetics* (pp. 67–79). Humana Press.

Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics, 19*(12), 1572–1574.

Shi, J., Blundell, T. L., & Mizuguchi, K. (2001). FUGUE: Sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *Journal of Molecular Biology, 310*(1), 243–257.

Shoemaker, J. S., Painter, I. S., & Weir, B. S. (1999). Bayesian statistics in genetics: A guide for the uninitiated. *Trends in Genetics, 15*(9), 354–358.

Stark, C., Breitkreutz, B. J., Reguly, T., Boucher, L., Breitkreutz, A., Livstone, M., et al. (2008). The BioGRID interaction database: 2008 update. *Nucleic Acids Research, 36*(suppl 1), D637–D640.

Steel, M., & Penny, D. (2000). Parsimony, likelihood, and the role of models in molecular phylogenetics. *Molecular Biology and Evolution, 17*(6), 839–850.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America, 102*(43), 15545–15550.

Tamura, K., Kumar, S., Nei, M., & Dudley, J. (2008). MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics, 9*(4), 299–306.

Thompson, J. D., Gibson, T., & Higgins, D. G. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Current Protocols in Bioinformatics,* 2–3.

Whelan, S., & Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Molecular Biology and Evolution, 18*(5), 691–699.

Xenarios, I., Salwinski, L., Duan, X. J., Higney, P., Kim, S. M., & Eisenberg, D. (2002). DIP, the database of interacting proteins: A research tool for studying cellular networks of protein interactions. *Nucleic Acids Research, 30*(1), 303–305.

Yang, Z. (1997). PAML: A program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences: CABIOS, 13*(5), 555–556.

Zmasek, C. M., & Eddy, S. R. (2002). RIO: Analyzing proteomes by automated phylogenomics using resampled inference of orthologs. *BMC Bioinformatics, 3*(1), 14.