Genetics and Regulation of Nitrogen Fixation in Free-Living Bacteria

Nitrogen Fixation: Origins, Applications, and Research Progress

VOLUME 2

Genetics and Regulation of Nitrogen Fixation in Free-Living Bacteria

Edited by

Werner Klipp

Lehrstuhl für Biologie der Mikroorganismen,
Fakultät für Biologie, Ruhr-Universität Bochum, Germany

Bernd Masepohl

Lehrstuhl für Biologie der Mikroorganismen,
Fakultät für Biologie, Ruhr-Universität Bochum, Germany

John R. Gallon

Biochemistry Research Group,
School of Biological Sciences, University of Wales, Swansea, U.K.

and

William E. Newton

Department of Biochemistry,
Virginia Polytechnic Institute & State University, Blacksburg, U.S.A.

KLUWER ACADEMIC PUBLISHERS
NEW YORK, BOSTON, DORDRECHT, LONDON, MOSCOW

eBook ISBN: 1-4020-2179-8
Print ISBN: 1-4020-2178-X

Created in the United States of America

Visit Springer's eBookstore at: http://ebooks.kluweronline.com
and the Springer Global Website Online at: http://www.springeronline.com

TABLE OF CONTENTS

SERIES PREFACE

Nitrogen Fixation: Origins, Applications, and Research Progress

Nitrogen fixation, along with photosynthesis as the energy supplier, is the basis of all life on Earth (and maybe elsewhere too!). Nitrogen fixation provides the basic component, fixed nitrogen as ammonia, of two major groups of macromolecules, namely nucleic acids and proteins. Fixed nitrogen is required for the N-containing heterocycles (or bases) that constitute the essential coding entities of deoxyribonucleic acids (DNA) and ribonucleic acids (RNA), which are responsible for the high-fidelity storage and transfer of genetic information, respectively. It is also required for the amino-acid residues of the proteins, which are encoded by the DNA and that actually do the work in living cells. At the turn of the millennium, it seemed to me that now was as good a time as any (and maybe better than most) to look back, particularly over the last 100 years or so, and ponder just what had been achieved. What is the state of our knowledge of nitrogen fixation, both biological and abiological? How has this knowledge been used and what are its impacts on humanity?

In an attempt to answer these questions and to capture the essence of our current knowledge, I devised a seven-volume series, which was designed to cover all aspects of nitrogen-fixation research. I then approached my long-time contact at Kluwer Academic Publishers, Ad Plaizier, with the idea. I had worked with Ad for many years on the publication of the Proceedings of most of the International Congresses on Nitrogen Fixation. My personal belief is that congresses, symposia, and workshops must not be closed shops and that those of us unable to attend should have access to the material presented. My solution is to capture the material in print in the form of proceedings. So it was quite natural for me to turn to the printed word for this detailed review of nitrogen fixation. Ad's immediate affirmation of the project encouraged me to share my initial design with many of my current co-editors and, with their assistance, to develop the detailed contents of each of the seven volumes and to enlist prospective authors for each chapter.

There are many ways in which the subject matter could be divided. Our decision was to break it down as follows: nitrogenases, commercial processes, and relevant chemical models; genetics and regulation; genomes and genomics; associative, endophytic, and cyanobacterial systems; actinorhizal associations; leguminous symbioses; and agriculture, forestry, ecology, and the environment. I feel very fortunate to have been able to recruit some outstanding researchers as co-editors for this project. My co-editors were Mike Dilworth, Claudine Elmerich, John Gallon, Euan James, Werner Klipp, Bernd Masepohl, Rafael Palacios, Katharina Pawlowski, Ray Richards, Barry Smith, Janet Sprent, and Dietrich Werner. They worked very hard and ably and were most willing to keep the volumes moving along reasonably close to our initial timetable. All have been a pleasure to work with and I thank them all for their support and unflagging interest.

Nitrogen-fixation research and its application to agriculture have been ongoing for many centuries – from even before it was recognized as nitrogen fixation. The Romans developed the crop-rotation system over 2000 years ago for maintaining and improving soil fertility with nitrogen-fixing legumes as an integral component. Even though crop rotation and the use of legumes was practiced widely but intermittently since then, it wasn't until 1800 years later that insight came as to how legumes produced their beneficial effect. Now, we know that bacteria are harbored within nodules on the legumes' roots and that they are responsible for fixing N_2 and providing these plants with much of the fixed nitrogen required for healthy growth. Because some of the fixed nitrogen remains in the unharvested parts of the crop, its release to the soil by mineralization of the residue explains the follow-up beneficial impact of legumes. With this realization, and over the next 100 years or so, commercial inoculants, which ensured successful bacterial nodulation of legume crops, became available. Then, in the early 1900's, abiological sources of fixed nitrogen were developed, most notable of these was the Haber-Bosch process. Because fixed nitrogen is almost always the limiting nutrient in agriculture, the resulting massive increase in synthetic fixed-nitrogen available for fertilizer has enabled the enormous increase in food production over the second half of the 20^{th} century, particularly when coupled with the new "green revolution" crop varieties. Never before in human history has the global population enjoyed such a substantial supply of food.

Unfortunately, this bright shiny coin has a slightly tarnished side! The abundance of nitrogen fertilizer has removed the necessity to plant forage legumes and to return animal manures to fields to replenish their fertility. The result is a continuing loss of soil organic matter, which decreases the soil's tilth, its water-holding capacity, and its ability to support microbial populations. Nowadays, farms do not operate as self-contained recycling units for crop nutrients; fertilizers are trucked in and meat and food crops are trucked out. And if it's not recycled, how do we dispose of all of the animal waste, which is rich in fixed nitrogen, coming from feedlots, broiler houses, and pig farms? And what is the environmental impact of its disposal? This problem is compounded by inappropriate agricultural practice in many countries, where the plentiful supply of cheap commercial nitrogen fertilizer, plus farm subsidies, has encouraged high (and increasing) application rates. In these circumstances, only about half (at best) of the applied nitrogen reaches the crop plant for which it was intended; the rest leaches and "runs off" into streams, rivers, lakes, and finally into coastal waters. The resulting eutrophication can be detrimental to marine life. If it encroaches on drinking-water supplies, a human health hazard is possible. Furthermore, oxidation of urea and ammonium fertilizers to nitrate progressively acidifies the soil – a major problem in many agricultural areas of the world. A related problem is the emission of nitrogen oxides (NO_x) from the soil by the action of microorganisms on the applied fertilizer and, if fertilizer is surface broadcast, a large proportion may be volatilized and lost as ammonia. For urea in rice paddies, an extreme example, as much as 50% is volatilized and lost to the atmosphere. And what goes up must come down; in the case of fertilizer nitrogen, it returns to Earth in the rain, often acidic in nature. This

uncontrolled deposition has unpredictable environmental effects, especially in pristine environments like forests, and may also affect biodiversity.

Some of these problems may be overcome by more efficient use of the applied fertilizer nitrogen. A tried and tested approach (that should be used more often) is to ensure that a balanced supply of nutrients (and not simply applying more and more) is applied at the right time (maybe in several separate applications) and in the correct place (under the soil surface and not broadcast). An entirely different approach that could slow the loss of fertilizer nitrogen is through the use of nitrification inhibitors, which would slow the rate of conversion of the applied ammonia into nitrate, and so decrease its loss through leaching. A third approach to ameliorating the problems outlined above is through the expanded use of biological nitrogen fixation. It's not likely that we shall soon have plants, which are capable of fixing N_2 without associated microbes, available for agricultural use. But the discovery of N_2-fixing endophytes within the tissues of our major crops, like rice, maize, and sugarcane, and their obvious benefit to the crop, shows that real progress is being made. Moreover, with new techniques and experimental approaches, such as those provided by the advent of genomics, we have reasons to renew our belief that both bacteria and plants may be engineered to improve biological nitrogen fixation, possibly through developing new symbiotic systems involving the major cereal and tuber crops.

In the meantime, the major impact might be through agricultural sustainability involving the wider use of legumes, reintroduction of crop-rotation cycles, and incorporation of crop residues into the soil. But even these practices will have to be performed judiciously because, if legumes are used only as cover crops and are not used for grazing, their growth could impact the amount of cultivatable land available for food crops. Even so, the dietary preferences of developed countries (who eats beans when steak is available?) and current agricultural practices make it unlikely that the fixed-nitrogen input by rhizobia in agricultural soils will change much in the near-term future. A significant positive input could accrue, however, from matching rhizobial strains more judiciously with their host legumes and from introducing "new" legume species, particularly into currently marginal land. In the longer term, it may be possible to engineer crops in general, but cereals in particular, to use the applied fertilizer more efficiently. That would be a giant step the right direction. We shall have to wait and see what the ingenuity of mankind can do when "the chips are down" as they will be sometime in the future as food security becomes a priority for many nations. At the moment, there is no doubt that commercially synthesized fertilizer nitrogen will continue to provide the key component for the protein required by the next generation or two.

So, even as we continue the discussion about the benefits, drawbacks, and likely outcomes of each of these approaches, including our hopes and fears for the future, the time has arrived to close this effort to delineate what we know about nitrogen fixation and what we have achieved with that knowledge. It now remains for me to thank personally all the authors for their interest and commitment to this project. Their efforts, massaged gently by the editorial team, have produced an indispensable reference work. The content is my responsibility and I apologize

upfront for any omissions and oversights. Even so, I remain confident that these volumes will serve well the many scientists researching nitrogen fixation and related fields, students considering the nitrogen-fixation challenge, and administrators wanting to either become acquainted with or remain current in this field. I also acknowledge the many scientists who were not direct contributors to this series of books, but whose contributions to the field are documented in their pages. It would be remiss of me not to acknowledge also the patience and assistance of the several members of the Kluwer staff who have assisted me along the way. Since my initial dealings with Ad Plaizier, I have had the pleasure of working with Arno Flier, Jacco Flipsen, Frans van Dunne, and Claire van Heukelom; all of whom provided encouragement and good advice – and there were times when I needed both!

It took more years than I care to remember from the first planning discussions with Ad Plaizier to the completion of the first volumes in this series. Although the editorial team shared some fun times and a sense of achievement as volumes were completed, we also had our darker moments. Two members of our editorial team died during this period. Both Werner Klipp (1953-2002) and John Gallon (1944-2003) had been working on Volume II of the series, *Genetics and Regulation of Nitrogen-Fixing Bacteria*, and that volume is dedicated to their memory. Other major contributors to the field were also lost in this time period: Barbara Burgess, whose influence reached beyond the nitrogenase arena into the field of iron-sulfur cluster biochemistry; Johanna Döbereiner, who was the discoverer and acknowledged leader in nitrogen-fixing associations with grasses; Lu Jiaxi, whose "string bag" model of the FeMo-cofactor prosthetic group of Mo-nitrogenase might well describe its mode of action; Nikolai L'vov, who was involved with the early studies of molybdenum-containing cofactors; Dick Miller, whose work produced new insights into MgATP binding to nitrogenase; Richard Pau, who influenced our understanding of alternative nitrogenases and how molybdenum is taken up and transported; and Dieter Sellmann, who was a synthetic inorganic chemistry with a deep interest in how N_2 is activated on metal sites. I hope these volumes will in some way help both preserve their scientific contributions and reflect their enthusiasm for science. I remember them all fondly.

Only the reactions and interest of you, the reader, will determine if we have been successful in capturing the essence and excitement of the many sterling achievements and exciting discoveries in the research and application efforts of our predecessors and current colleagues over the past 150 years or so. I sincerely hope you enjoy reading these volumes as much as I've enjoyed producing them.

William E. Newton
Blacksburg, February 2004

PREFACE

Genetics and Regulation of Nitrogen-Fixing Bacteria

This book is the second volume of a seven-volume series, which covers all fields of research related to nitrogen fixation - from basic studies through applied aspects to environmental impacts. Volume II provides a comprehensive and detailed source of information concerning the genetics and regulation of biological nitrogen fixation in free-living prokaryotes. This preface attempts to provide the reader with some insight into how this volume originated, how it was planned, and then how it developed over the several years of its production.

Once the editorial team was established, the first job was to decide which of the many free-living diazotrophs that have been subjected to genetic analysis should be included in this volume. Would we need to develop specific criteria for selection or would the organisms, in effect, select themselves? Of course, *Klebsiella pneumoniae* and *Azotobacter vinelandii*, which have served (and still serve) as the main model organisms for the genetic analysis of diazotrophy, plus some of the other bacteria described in this volume, did indeed select themselves. However, there was considerable discussion surrounding well-characterized fixing species, like *Azorhizobium caulinodans* and *Herbaspirillum seropedicae*, both of which are able to fix atmospheric N_2 under free-living conditions. Was this volume the right place for them? If they were omitted here, would it compromise the volume as a major reference work? After discussions both among ourselves and with the editorial teams of other volumes, it was finally agreed that bacteria such as these belong elsewhere. We decided, for better or worse, that their ability to fix N_2 in either a symbiotic or close associative interaction with a plant host was the significant differentiating factor and so, they are described in other volumes of this series that deal specifically with these interactions and the involved partners.

Similar concerns arose with, for example, coverage of the superoxide-dependent nitrogenase system of *Streptomyces thermoautotrophicus*. How should this system be treated? The decision again was that this unique nitrogenase should be described in volume I of this series, *Catalysts for Nitrogen Fixation: Nitrogenases, Relevant Chemical Models and Commercial Processes*, which deals with the structural and functional aspects of nitrogenases and related non-biological systems, because there is only very limited knowledge of the genetics and regulation in this system.

After the questions concerning which organisms should be included were resolved, there was still the question of what more general topics must be covered. As you will see, we decided to include, in addition to the individual chapters representing an organism-based view of selected well-characterized diazotrophic proteobacteria, cyanobacteria, Gram-positive clostridia, and archaea, the more general cross-organismic themes dealing with different regulatory aspects, electron transport to nitrogenase, and molybdenum metabolism. In all the chapters, whereever appropriate, historical aspects have been included to give the reader a

sense of where things started and how much has been achieved, especially in the last 25 years or so.

The chapters are ordered with respect to the two themes mentioned above; a longer organism-based section and a shorter more general section. We are well aware that there are many possibilities for ordering the themes and the individual chapters of this volume and, no matter which is chosen, there are always many arguments either for or against any particular organization.

Our collective wisdom was to start with the organism-based theme and with chapters on the two model organisms. So first, the historical aspects of the genetics and regulation of nitrogen fixation are covered in Chapter 1, which is authored by Ray Dixon. Because of its close relationship to *Escherichia coli, K. pneumoniae* was the organism of choice for the initial genetics and regulation studies of nitrogen fixation. It turned out to be a very good choice for another, completely unexpected, reason; *K. pneumoniae* harbors only one nitrogenase and this is the classical molybdenum-containing nitrogenase. In contrast, some *Azotobacter* species, the subject of Chapter 2, which is co-authored by Christina Kennedy and Paul Bishop, including *A. vinelandii*, harbor as many as two alternative nitrogenases in addition to the classical molybdenum-nitrogenase. These two alternative nitrogenases, which contain either vanadium or iron in place of molybdenum, are regulated differently to the Mo-nitrogenase and so, of course, complicate the genetics and regulation within these organisms. The volume continues with two shorter papers on distantly related groups of prokaryotes outside the proteobacteria. These are the clostridia (Chapter 3, authored by J.-S. Chen) and archaea (Chapter 4, authored by John Leigh). Many of the organisms in these groups have been recalcitrant to genetic studies and much less is known about how nitrogen fixation is regulated. These chapters are followed by chapters about three different groups of photosynthetic bacteria, namely the heterocyst-forming cyanobacteria (Chapter 5, authored by Terry Thiel), the non-heterocystous cyanobacteria (Chapter 6, authored by John Gallon), and the purple bacteria (Chapter 7, co-authored by Bernd Masepohl, Thomas Drepper, and Werner Klipp).

The move to the more general cross-organismic theme starts with post-translational regulation of nitrogenase in photosynthetic bacteria (Chapter 8, co-authored by Stefan Nordlund and Paul Ludden), which allows a seamless transition from the chapter on purple bacteria to the review on regulation of nitrogen fixation in free-living diazotrophs (Chapter 9, authored by Mike Merrick). Two more chapters, one on molybdenum metabolism (Chapter 10, authored by Richard Pau) and the other on electron transport to nitrogenase (Chapter 11, authored by Kaz Saeki) complete the general overview section. The volume ends with a chapter (Chapter 12, co-authored by John Gallon and Bernd Masepohl) that considers the prospects for, and provides an outlook on, the future of this area of scientific endeavor.

We wish to thank all the authors of this volume for their enormous efforts to make it an indispensable reference work for all scientists working in the field, for those administrators with authority and responsibility in this arena, and for students who are brave enough to want to enter this challenging area of research. There are many other experts in the field who could have made contributions to this volume,

but who were not asked to do so. We wish to make it crystal clear that our choices in no way reflect on them and their abilities. In fact, this volume could not have been produced without their contributions to the field, many of which are incorporated into these pages. In addition, we proffer our apologies in advance if any topics of interest are omitted. At this point, we would like to draw special attention to another volume in this series, which is entitled *Genomes and Genomics of Nitrogen-Fixing Organisms*. It deals with the whole genome of the many nitrogen-fixing bacteria (including species described in this volume), whose genomes have been completely sequenced. There, a considerable amount of information is gathered, much of it closely related to the topics of this volume, and it is also the volume in which the evolutionary aspects of nitrogen fixation are considered.

It took more than two years from the first phase of planning until the completion of this volume. Both the fun and the sense of achievement that should have accrued during this time were overshadowed by the deaths of three contributors to this volume. Werner Klipp, who helped start this project, died in May 2002 towards the end of the major planning phase. And, as if this wasn't enough, John Gallon, who had been recruited and was actively working on this volume after Werner´s death, died in August 2003 during the final preparation of the individual chapters. Then, in February 2004 after completion of the volume, Richard Pau passed away.

With great sadness, some months ago, we decided to dedicate this volume to our two co-editors and good friends, who were lost on the journey to complete this project. And now, as we look back, we fondly remember all three of them, not just for their keen scientific insight and research prowess, but for their enthusiasm for science, for the joy they took from scientific discovery, and for sharing the fun with us along the way.

Bernd Masepohl
Bochum, February 2004

William E. Newton
Blacksburg, February 2004

LIST OF CONTRIBUTORS

Paul E. BISHOP
Department of Microbiology, North Carolina State University and USDA, Raleigh, NC 27695, U.S.A.
Email: paul_bishop@ncsu.edu

Jiann-Shin CHEN
Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, U.S.A.
Email: chenjs@vt.edu

Ray DIXON
Department of Molecular Microbiology, John Innes Centre, Norwich NR2 3DR, U.K.
Email: ray.dixon@bbsrc.ac.uk

Thomas DREPPER
Ruhr-Universität Bochum, Fakultät für Biologie, Lehrstuhl für Biologie der Mikroorganismen, D-44780 Bochum, Germany.
Email: thomas.drepper@ruhr-uni-bochum.de

John R. GALLON
Biochemistry Research Group, School of Biological Sciences, University of Wales Swansea, Singleton Park, Swansea, SA2 8PP, U.K.

Christina KENNEDY
Department of Plant Pathology, University of Arizona, Tucson, AZ 85721, U.S.A.
Email: ckennedy@u.arizona.edu

Werner KLIPP
Ruhr-Universität Bochum, Fakultät für Biologie, Lehrstuhl für Biologie der Mikroorganismen, D-44780 Bochum, Germany

John A. LEIGH
Department of Microbiology, University of Washington, Seattle, WA 98195, U.S.A.
Email: leighj@u.washington.edu

Paul W. LUDDEN
Department of Plant and Microbial Biology, 111 Koshland Hall, University of California, Berkeley, CA 94720-3102, U.S.A.
Email: pludden@nature.berkeley.edu

Bernd MASEPOHL
Ruhr-Universität Bochum, Fakultät für Biologie, Lehrstuhl für Biologie der Mikroorganismen, D-44780 Bochum, Germany.
Email: bernd.masepohl@ruhr-uni-bochum.de

Mike MERRICK
Department of Molecular Microbiology, John Innes Centre, Norwich, NR4 7UH, U.K.
Email: mike.merrick@bbsrc.ac.uk

William E. NEWTON
Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, U.S.A.
Email: wenewton@vt.edu

Stefan NORDLUND
Department of Biochemistry and Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden.
Email: stefan@dbb.su.se

Richard N. PAU
Department of Biochemistry and Molecular Biology, The University of Melbourne, Parkville 3010, Australia.
Email: r.pau@unimelb.edu.au

Kazuhiko SAEKI
Department of Biology, Graduate School of Science, Osaka University, 1-1 Machikaneyama, Toyonaka, Osaka 560-0043, Japan.
Email:ksaeki@bio.sci.osaka-u.ac.jp

Teresa THIEL
Department of Biology, University of Missouri-St. Louis, 8801 Natural Bridge Rd., St. Louis, MO 63121, U.S.A.
Email: thiel@umsl.edu

John Gallon
(1944-2003)

Werner Klipp
(1953-2002)

We dedicate this volume to two of our colleagues, Werner Klipp and John Gallon, both of whom died during the period when we were preparing this work for publication. Both were well-known and well-respected members of the international research community with interests that ranged widely across many aspects of science related to nitrogen-fixation research. Werner's interests covered many of the molecular aspects related to both free-living and symbiotic bacteria, including molybdenum metabolism, but he was especially interested in genetics and regulation. His enthusiasm for his work was obvious to those of us who were fortunate enough to know him well and was matched only by his brilliant rapid-fire seminars. Without his input, the field would be poorer by far with respect to our detailed knowledge of nitrogen fixation, especially in photosynthetic bacteria. John's interests encompassed the whole organism, particularly the non-heterocystous cyanobacteria, rather than individual cell components. His "bigger picture" interests led him to research areas where complex interactions are the norm. His work at the interface of the organism with its environment has helped us in our understanding of how the environment impacts the organism's ability to survive and grow and, just as importantly, how organisms impact their environment. Both Werner and John enjoyed their science immensely; discussions with either of them were always great fun. Many of us will remember them both as first-class researchers, constructive collaborators, fine teachers, and good friends.

Chapter 1

HISTORICAL PERSPECTIVE – DEVELOPMENT OF *NIF* GENETICS AND REGULATION IN *KLEBSIELLA PNEUMONIAE*

R. DIXON
Department of Molecular Microbiology, John Innes Centre, Norwich NR2 3DR, UK

"History never looks like history when you are living through it."
John W. Gardner

1. INTRODUCTION

By the end of the 1960's, research on biological nitrogen fixation had progressed to the extent that genetic analysis of this process and its regulation had clearly become feasible. Biochemical studies during this decade had established methods for the purification of nitrogenase component proteins and the advent of the acetylene reduction test had revolutionised measurement of nitrogen fixation both *in vitro* and *in vivo*. Pathways for ammonia assimilation under nitrogen-limited conditions had been characterised and physiological studies had suggested a role for fixed nitrogen in "repression" of nitrogenase synthesis.

Mutants of *Azotobacter vinelandii* incapable of fixing N_2 were first isolated as early as 1950, but such mutants were not particularly useful at that time, because there was no genetic method available to map the mutations, nor was it possible to examine the properties of mutant proteins at the biochemical level. Clearly, genetic analysis would require the development of gene transfer systems in nitrogen-fixing bacteria. The rapid advances in microbial genetics during the 1960's provided the impetus to investigate the genetics of nitrogen fixation and, in the early 1970's, *Klebsiella pneumoniae* became the model organism for genetic analysis of diazotrophy. In considering the development of *nif* genetics from a historical standpoint, the 1970's therefore provide an obvious starting-point for this review.

W. Klipp, B. Masepohl, J. R. Gallon and W. E. Newton (eds.), Genetics and Regulation of Nitrogen Fixation in Free-Living Bacteria, 1-25.

However, in writing this Chapter, it has been more difficult to decide where the story should end, as past developments clearly form a continuum with current research. I have arbitrarily chosen the mid-1990's as a closing point, primarily because the most recent advances in the field will be covered by other authors in this volume.

2. THE EARLY YEARS

2.1. nif Gene Transfer

In 1970, two laboratories independently initiated genetic studies on nitrogen fixation in the gram-negative bacterium *Klebsiella pneumoniae* strain M5a1. This organism was chosen primarily because it was an enteric diazotroph related to *Escherichia coli* and *Salmonella typhimurium*, both of which had been extensively studied at the genetic level. This diazotrophic non-mucoid *K. pneumoniae* strain had also been used for biochemical and physiological studies of nitrogen fixation. Genetic recombination either by transduction or conjugation had been reported in some *Klebsiella* strains but not in those which were competent to fix N_2. Stanley Streicher, a graduate student working in Ray Valentine's laboratory at La Jolla isolated *nif* mutants of *K. pneumoniae* M5a1 unable to grow on N_2 as sole nitrogen source and showed that these mutants were defective in catalysing acetylene reduction (Streicher *et al.*, 1971). Although strain M5al was not sensitive to bacteriophage active on other *Klebsiella* strains, Streicher *et al*. found that this diazotrophic strain was sensitive to phage P1, which had been used previously as a tool for generalised transduction in *E. coli*. P1 lysates prepared on wild-type strain M5a1, transduced the *nif* $^-$ mutants to a Nif$^+$ phenotype, allowing the transductants to regain the ability to grow on N_2 as sole nitrogen source. A series of two-point transductional crosses with 30 mutants revealed that *nif* genes were located in one region of the chromosome, close to the histidine (*his*) biosynthetic operon (Streicher *et al.*, 1971).

In contrast to the transduction approach, I used conjugation to establish *nif* gene transfer in *K. pneumoniae* M5a1 while working as a graduate student in John Postgate's laboratory at Sussex (Dixon and Postgate, 1971). The self-transmissible R plasmid, R144drd3, was observed to promote chromosome mobilisation in *K. pneumoniae* and, when transferred from a wild-type donor strain to a *nif* $^-$ recipient strain, gave rise to Nif$^+$ recombinants at frequencies around 10^{-5}. These studies also demonstrated that *nif* mutations were located close to the *his* operon in *K. pneumoniae*. In order to increase the frequency of conjugal gene transfer, we irradiated the M5a1 (R144drd3) strain with UV light to obtain an Hfr-like strain, which was subsequently shown to mobilise the *his* region of the chromosome in a polarised fashion (Dixon *et al.*, 1975; Dixon and Postgate, 1972). This high frequency donor strain was used to attempt conjugal transfer of the *K. pneumoniae his* region to *E. coli.* Remarkably, some of the His$^+$ *E. coli* transconjugants gained the ability to fix N_2 and thus the first genetically engineered diazotroph had been created (Dixon and Postgate, 1972).

At the time, the genetic transfer of nitrogen fixation to this extensively studied model prokaryote was regarded as an exceptional breakthrough. It indicated that all the genes required to enable *E. coli* to fix N_2 were located close to the histidine biosynthetic operon in *K. pneumoniae* and that rapid genetic analysis of nitrogen fixation would be facilitated by the powerful genetic tools available in *E.coli*. It also led to speculation concerning future prospects for genetic engineering of nitrogen fixation and the ultimate goal of generating autonomous nitrogen-fixing plants (Shanmugam and Valentine, 1975b; Streicher *et al.*, 1972). During this period, genetic transfer of nitrogen fixation from *Rhizobium trifolii* to a non-nitrogen fixing strain of *K. aerogenes* was also reported (Dunican and Tierney, 1974), although this finding has not been confirmed by other laboratories.

Analysis of the nitrogen-fixing *E. coli* transconjugants revealed that *K. pneumoniae nif* genes were either integrated into the *E. coli* chromosome or were located on autonomously replicating plasmids (Cannon *et al.*, 1974a, 1974b). Selection for recombinant plasmids *in vivo* led to the construction of a wide host range plasmid of the P incompatibility group carrying the *his-nif* region of *K. pneumoniae*, which was transmissible to *Agrobacterium tumefaciens* and *Rhizobium meliloti* (Dixon *et al.*, 1976). The most useful property of this plasmid, however, was not its wide host range, but its stability in *K. pneumoniae* and it was used extensively for subsequent complementation analysis and cloning of the *nif* gene cluster (see below).

2.2. Regulation by Ammonium

Tempest's discovery of a new route for ammonia assimilation, involving the enzymes glutamate synthase (GOGAT) and glutamine synthetase (GS) in *Klebsiella aerogenes* (Tempest *et al.*, 1970), was soon extended to nitrogen-fixing bacteria. Mutants of *K. pneumoniae* M5a1 defective in glutamate synthase (*asm* mutants) were unable to assimilate nitrogen under nitrogen-fixing conditions but were able to grow in media containing ammonia as the nitrogen source (Nagatani *et al.*, 1971). These observations strongly suggested that the GS-GOGAT pathway was essential for ammonia assimilation under diazotrophic conditions and also focussed attention on the potential role of GS in regulating nitrogenase synthesis in response to ammonium. Based upon studies on a variety of mutants with lesions in *glnA*, the structural gene for GS, Magasanik and his colleagues had proposed that GS was not only a key enzyme in ammonia assimilation, but also played a key role in regulating transcription of nitrogen-regulated operons (Magasanik *et al.*, 1974). This proposal was backed up by the observation that purified GS stimulated transcription of the histidine utilisation operon *in vitro* (Tyler *et al.*, 1974). The proposed regulatory role for GS in nitrogen regulation was supported by studies on ammonium regulation of nitrogen fixation. Independent studies by a graduate student (Roy Tubb) working in Postgate's laboratory and Stanley Streicher, who had moved from Valentine's group to MIT, led to similar conclusions. Mutants defective in glutamine synthetase showed no detectable nitrogenase activity, but mutants with constitutive levels of GS activity synthesised nitrogenase in the presence of

ammonium (Shanmugam *et al.*, 1975; Streicher *et al.*, 1974; Tubb, 1974). These findings led to the hypothesis that GS positively controlled *nif* transcription as proposed for other nitrogen-regulated genes (Magasanik *et al.*, 1974; Prival *et al.*, 1973). This hypothesis survived in the literature for ten years until it was eventually realised that nitrogen regulation was mediated not by GS but by the products of regulatory genes located in the same operon as *glnA* itself (see below).

The early observations on nitrogen assimilation in diazotrophic bacteria were, however, further exploited to isolate mutant *K. pneumoniae* strains that excreted ammonia when grown under nitrogen-fixing conditions (Shanmugam and Valentine, 1975a). Although this was heralded as a potential route for microbial production of ammonium-based fertiliser, the energetic cost of biological nitrogen fixation precluded any exploitation of this finding on a large scale.

3. DEFINING THE *K. PNEUMONIAE NIF* GENES

Biochemical studies on nitrogenase had shown that the enzyme was comprised of two component proteins, suggesting that only a few genes would be required for nitrogenase biosynthesis. In the latter half of the 1970s, several laboratories embarked on identifying *nif* genes and determining their operon structure. The major competition was between Winston Brill's group in Madison and the Sussex group. Brill's group had initially mapped *K. pneumoniae nif* mutants by three-factor transduction crosses with phage P1 using *his* as an outside marker (St John *et al.*, 1975). This approach was followed by complementation analysis of *nif* mutants with plasmid pRD1 and further P1 transduction experiments, which suggested that the *nif* region was comprised of at least seven cistrons (Dixon *et al*., 1977), divided into a proximal *his* cluster, containing *nifB*, *nifA* (*nifL*), *nifF*, and a more distal cluster, comprising *nifE*, *nifK*, *nifD* and *nifH* (Kennedy, 1977). Brill's group exploited phage Mu to isolate a series of *nif* deletion mutants both in the chromosome and on a derivative of pRD1 (Bachhuber *et al.*, 1976; Macneil *et al.*, 1978a). These deletions were used to map an extensive collection of point mutants that closed the gap between the proximal and distal *nif* clusters identified by (Kennedy, 1977) and identified a total of 14 *nif* genes designated as *nifQ, nifB, nifA, nifL, nifF, nifM, nifV, nifS, nifN, nifE, nifK, nifD, nifH and nifJ* (Macneil *et al.*, 1978b). A group in China led by Shen also mapped a series of *nif* mutations to close this gap (Hsueh *et al.*, 1977).

Meanwhile, the Sussex group had identified four new *nif* genes, which they designated as *nifM*, *nifN, nifI* and *nifJ* (Merrick *et al.*, 1978). This nomenclature caused some confusion at the time, because *nifN* was equivalent to the gene designated *nifS* by Brill's group and *nifI* was equivalent to *nifN*. This conflict was resolved after a heated discussion at the 3rd International Symposium on Nitrogen Fixation in Madison, Wisconsin, in June 1978, whereby the Brill nomenclature was accepted. Around this time, Claudine Elmerich's laboratory at the Institut Pasteur in Paris isolated Mu insertions in *nif* genes. By examining the transcriptional polarity of Mu or other transposon-induced mutations in complementation tests, all three laboratories suggested a similar operon structure for the *nif* gene cluster

(Elmerich *et al.*, 1978; Macneil *et al.*, 1978b; Merrick *et al.*, 1978). Further collaboration between the Sussex and Paris groups led to a more detailed fine-structure map which included a newly identified gene designated *nifU,* located between *nifS* and *nifN* (Merrick *et al.*, 1980).

Although *nifH*, *nifD* and *nifK* were soon identified as the structural genes for nitrogenase, the complexity of the *nif* gene cluster came as a surprise to the nitrogen fixation community and the immediate challenge was to determine the functions of each of the gene products. Gary Roberts working in Brill's laboratory identified *nif* polypeptides on two-dimensional gels and examined nitrogenase activity in mutant crude extracts, following *in vitro* complementation with purified Fe protein, MoFe protein or iron-molybdenum cofactor (FeMoco). Fifteen *nif*-encoded polypetides were identified on gels. The *nifB, nifE* and *nifN* polypeptides were shown to be required for FeMoco synthesis and the *nifM* and *nifS* genes for the synthesis of active Fe protein (Roberts and Brill, 1980; Roberts *et al.*, 1978). Although *nifF* and *nifJ* were both required for nitrogenase activity *in vivo*, crude extracts of these mutants were active *in vitro,* implying a role for these proteins in electron transport to nitrogenase (Hill and Kavanagh, 1980; Nieva-Gomez *et al*., 1980; Roberts *et al*., 1978). Sequencing of *nif* genes provided useful clues to the possible functions of other *nif*-encoded polypeptides (see below).

4. THE RECOMBINANT DNA ERA

Several nitrogen-fixation scientists were invited to the Asilomar Conference in 1975, which led to the establishment of guidelines for working with the newly discovered recombinant DNA technology. Once the guidelines were established, it was not long before the cloning of *nif* DNA began. Frank Cannon at Sussex, working in collaboration with Fred Ausubel's laboratory at Harvard, isolated DNA from the *nif* plasmid pRD1 and after partial digestion with EcoRI, inserted fragments into the amplifiable plasmid pMB9. Ligated DNA was transformed into a restriction-deficient *K. pneumoniae* strain containing mutations in *hisD* and *nifB*. His^+ transformants were screened for those which complemented the *nifB* mutation. One plasmid, designated as pCRA37, complemented *nifB* and *nifF* mutations but not *nifD* or *nifH* mutants; the physical map of this recombinant plasmid correlated well with genetic map derived by P1 co-transductional analysis (Cannon *et al*., 1977). The remaining *nif* genes were cloned in subsequent stages, culminating in the construction of multicopy plasmids carrying the entire *nif* region (Cannon *et al*., 1979; Pühler *et al*., 1979; Reidel *et al*., 1979). The availability of *K. pneumoniae nif* DNA fragments immediately facilitated the cloning of *nif* genes from other diazotrophs using Southern hybridisation (Ruvkun and Ausubel, 1980). The following year, the entire *K. pneumoniae nif* gene cluster was stably integrated into the genome of yeast following transformation with an *E. coli*-yeast shuttle vector (Zamir *et al*., 1981). Although this experiment was an interesting feat of genetic engineering, it had little practical relevance because no attempt was made to express the prokaryotic nif genes from yeast promoters. Expression of *nifH*, *nifD*, and *nifK*

was subsequently achieved in yeast but the nitrogenase component proteins were apparently inactive in this host (Berman *et al*., 1985; Holland *et al*., 1987).

The advent of the recombinant DNA era and the possibilities for exploiting the technology to improve biological nitrogen fixation generated considerable interest during the late 1970´s and a conference on "Genetic Engineering for Nitrogen Fixation" sponsored by the US National Science Foundation was held at the Brookhaven National Laboratory in 1977 to discuss these issues. Although the potential for transferring nitrogen-fixation genes from prokaryotes directly into plants was considered at this meeting, in retrospect this conference was far too premature because plant transformation had not been developed. It is interesting, however, to consider the motivation for engineering biological nitrogen fixation at that time, because there was rising concern about the energy costs of fertiliser production. One projection suggested: "assuming a steady supply of natural gas at current levels, in the year 2000, 10% of the entire U.S. supply (2 trillion cubic feet) will be consumed by fertiliser factories" (Ausubel, 1977). This was, of course, a gross over-estimation because current industrial fertiliser production accounts for only around 1-2% of worldwide energy consumption.

The new chemical and dideoxy methods for sequencing DNA (Maxam and Gilbert, 1977; Sanger *et al.*, 1977) were soon applied to determine the nucleotide sequences of *nif* genes. The first sequence to be published was that of *nifH,* from the cyanobacterium *Anabaena* 7120 by Bob Haselkorn's laboratory in Chicago (Mevarech *et al.*, 1980). This sequence was closely followed by that of *K. pneumoniae nifH* and *nifD* (Scott *et al.*, 1981; Sundaresan and Ausubel, 1981). Further sequence analysis revealed five new *nif* genes *nifT*, *nifW, nifX, nifY*, *nifZ* and the complete DNA sequence of the entire 24.2 Kb *K. pneumoniae nif* gene cluster was completed in 1988 (Arnold *et al.*, 1988). The availability of the sequence data provided an extremely useful platform to establish the functions of *nif* gene products, for example, the role of *nifM* in Fe protein biosynthesis (Paul and Merrick, 1987, 1989) and the *nifU* and *nifS* genes in iron-sulphur cluster biosynthesis (Beynon *et al.*, 1987; Dean and Jacobson, 1992). A description of the role of *nif* genes in nitrogenase biosynthesis and function is given in Table 1.

5. *NIF* GENE REGULATION

5.1. Identification of the Major Regulatory Genes

The observation that nitrogenase component proteins were absent in *nifA* mutants led to the suggestion that the *nifA* gene product was a positive activator of *nif* gene expression (Dixon *et al.*, 1977). A new gene, *nifL*, upstream of *nifA*, was defined on the basis of transductional mapping of a mutant allele, *nifL2265*, which was partially transdominant to the wild-type allele (Kennedy, 1977). Mu-induced insertion mutations in *nifL* were polar on *nifA*, but this polarity could be relieved by suppressor mutations, suggesting that *nifL* is not essential for nitrogen fixation O_2 (Macneil *et al.*, 1978b). Although repression of nitrogenase biosynthesis by ammonia was a recognised phenomenon at that time (section 2.2.), repression in

Table 1. Proposed functions of K. pneumoniae *nif genes*

Gene	*Product and/or known function*	*References*
nifJ	Pyruvate oxido-reductase; couples pyruvate oxidation to reduction of the *nifF* product	(Hill and Kavanagh, 1980; Nieva-Gomez *et al.*, 1980; Shah *et al.*, 1983)
nifH	Fe protein subunit	(Dean and Jacobson, 1992)
nifDK	MoFe protein α and β subunits	(Dean and Jacobson, 1992)
nifT	Function unknown; not essential for nitrogen fixation	(Simon *et al.*, 1996)
nifY	Associates with MoFe protein and dissociates upon FeMo cofactor insertion	(Homer *et al.*, 1993)
nifEN	Required for FeMo cofactor biosynthesis	(Allen *et al.*, 1994; Dean and Jacobson, 1992)
nifX	Not essential for nitrogen fixation; required for FeMo cofactor biosynthesis	(Shah *et al.*, 1999)
nifU	Fe-S cluster biosynthesis	(Yuvaniyama *et al.*, 2000)
nifS	Fe-S cluster biosynthesis	(Zheng *et al.*, 1993)
nifV	Encodes a homocitrate synthase. Homocitrate is an organic component of FeMo cofactor	(Hawkes *et al.*, 1984; Hoover *et al.*, 1987; McLean and Dixon, 1981)
nifW	Function unknown; interacts with the MoFe protein	(Kim and Burgess, 1996; Paul and Merrick, 1989)
nifZ	Function unknown; required for full activity of the MoFe protein	(Paul and Merrick, 1989)
nifM	Required for Fe protein maturation. Putative peptidyl-prolyl *cis/trans* isomerase	(Dean and Jacobson, 1992; Gavini and Pulukat, 2002)
nifF	Flavodoxin required for electron transfer to the Fe protein	(Deistung *et al.*, 1985; Thorneley *et al.*, 1992)
nifL	Negative regulatory protein	(Dixon, 1998)
nifA	Positive regulator of *nif* transcription	(Dixon, 1998)
nifB	Required for FeMo cofactor biosythesis	(Allen *et al.*, 1995; Shah *et al.*, 1994)
nifQ	Incorporation of Mo into FeMo cofactor	(Imperial *et al.*, 1984)

response to O_2 was not firmly established until 1978, when it was demonstrated that repression could be distinguished mechanistically from regulation by ammonium (Eady *et al.*, 1978). In 1979, an elegant technique had been developed to fuse any promoter to the *E. coli* lactose operon (Casadaban and Cohen, 1979) and this was soon exploited to isolate *nif-lac* fusions (Dixon *et al.*, 1980; Macneil *et al.*, 1981). These studies revealed that all *nif* promoters respond to repression by ammonium and all except the *nifLA* promoter were strongly repressed by O_2. Similar findings were observed subsequently when rates of *nif* transcription were measured (Cannon *et al.*, 1985; Collins and Brill, 1985). The *nifA* gene product was shown to act *in trans* as a positive activator of *nif* gene expression, but was not required to activate the *nifLA* promoter (Dixon *et al.*, 1980). A strange, and as yet unexplained phenomenon, is that Nif^+ strains give rise to a purple colour on nitrogen-free medium containing 6-cyanopurine and this phenotype requires the *nifA* gene product (Macneil and Brill, 1978). A few of the suppressors of polar insertion mutations in *nifL* were able to synthesise nitrogenase in the presence of ammonium and were purple on 6-cyanopurine plates (Macneil *et al.*, 1981). These mutations, which were closely linked to *nifL*, also suppressed the Nif^- phenotype of *gln* mutants, leading to the hypothesis that *gln*-encoded factors might activate the *nifLA* promoter and that the *nifA* product specifically activates transcription of the other *nif* operons (Dixon *et al.*, 1980; Macneil *et al.*, 1981).

In the late 1970´s, the hypothesis that GS, in its non-adenlylylated form, was a positive regulator of nitrogen-regulated genes, including the *nif* operons, was beginning to wane, although there was some evidence that purified GS bound specifically to *nif* promoters (Janssen *et al.*, 1980). However, the simple model for GS as a nitrogen-responsive transcriptional activator was clearly incomplete because three new genes influencing nitrogen regulation had been discovered; *ntrA* (formerly *glnF*) unlinked to *glnA* (Garcia *et al.*, 1977) and *ntrB* (*glnL*) and *ntrC (glnG)* (McFarland *et al.*, 1981; Pahel and Tyler, 1979), which where shown to be in the same operon as *glnA* (Macneil *et al.*, 1982a; Pahel *et al.*, 1982). The earlier evidence that *glnA*-linked mutations influenced regulation of nitrogen-regulated genes could be explained on the basis of their polar effects on *ntrB* and *ntrC*. The first evidence for the role of these new genes in *nif* gene regulation was obtained by Leonardo and Goldberg (1980), who demonstrated that *ntrA* and *ntrC* mutants of *K. pneumoniae* were unable to synthesise nitrogenase. Subsequent studies in other laboratories confirmed these conclusions (de Bruijn and Ausubel, 1981; Espin *et al.*, 1982; Espin *et al.*, 1981) and at this stage, it became obvious that both *ntrC* and *ntrA* were required for positive control of *nif* transcription. Constitutive expression of *nifA* led to activation of *nif* transcription in the absence of *ntrC*, suggesting that *ntrC* is required to activate the *nifLA* promoter (Buchanan-Wollaston *et al.*, 1981a). However, constitutive expression of *nifA* did not by-pass the requirement for *ntrA*, suggesting that the *ntrA* product had a more general role in *nif* gene activation (Merrick, 1983; Ow and Ausubel, 1983; Sibold and Elmerich, 1982). These findings were elaborated into a cascade model for *nif* gene activation in which NtrC, acting in concert with NtrA, activated transcription of the *nifLA* operon, giving rise

to the expression of NifA, which activated transcription of the other six *nif* operons, also in concert with NtrA.

While attention was focussed on the *ntrC* and *nifA* gene products as positive activators of *nif* transcription, there was increasing evidence that *nifL* was involved in negative control. Several regulatory mutants mapping in *nifL* enabled *nif* transcription in the presence of O_2 but not in the presence of fixed nitrogen (Hill *et al.*, 1981). Subsequently, *nifL* was shown also to mediate repression in response to fixed nitrogen (Merrick *et al.*, 1982). When present on a multicopy plasmid, *nifL* inhibited *nif* transcription, even in the absence of O_2 and fixed nitrogen, suggesting that over-expression of this protein might titrate out factors required to maintain it in an inactive form (Buchanan-Wollaston *et al.*, 1981b; Reidel *et al.*, 1983). Based on stability analysis of *nif* mRNA, it was also proposed that the *nifL* product might have a role in post-transcriptional control of *nif* genes by destabilising *nif* mRNA in response to O_2 and fixed nitrogen (Collins *et al.*, 1986). However, this model has since been disproved and current analysis indicates that the stability of *nif* structural mRNA is associated with nitrogenase activity (Simon *et al.*, 1999).

The newly-discovered *ntrB* gene provided a focus for analysis as a potential negative regulator of *ntrC*. Several laboratories had isolated mutations in this gene that gave rise to constitutive expression of GS and other operons normally subject to repression by ammonium (Chen *et al.*, 1982; Macneil *et al.*, 1982b; McFarland *et al.*, 1981). Mutations in *ntrB* also gave rise to *nif* transcription in the presence of ammonium (Ow and Ausubel, 1983). These results suggested that *ntrB,* like *nifL*, had a negative function when ammonium was present. It was also thought that *nif* transcription could be subject to the stringent response under nitrogen-limiting conditions because *relA* mutants, which are unable to accumulate ppGpp, gave low rates of *nif* expression under nitrogen-starved conditions (Riesenberg *et al.*, 1982). However, it was subsequently found that there was no correlation between the levels of ppGpp and the extent of *nif* depression (Nair and Eady, 1984).

5.2. Sequencing of nif *Promoters and Regulatory Genes*

Determination of transcription start-sites and sequencing of the *nif* promoters revealed some surprises. Analysis of the *nifLA* promoter demonstrated that it had no typical –35 region and, in agreement with this finding, some positive control by NtrC was maintained even in deletions extending to –28 (Drummond *et al.*, 1983). Analysis of promoters activated either by NtrC or NifA revealed a heptameric consensus sequence, TTTGCA, in the –15 region, which was proposed to be a binding sequence for transcriptional activation (Ow *et al.*, 1983). In a seminal paper, sequencing of five *nif* promoters revealed a characteristic primary structure with the consensus CTGG at –24 and TTGCA at –12 (Beynon *et al.*, 1983). The significance of this unique consensus sequence was not then fully realised, although it was postulated that this might provide a recognition sequence for a modified form of RNA polymerase containing the *ntrA* product, which might act as a novel sigma factor (de Bruijn and Ausubel, 1983). Footprinting experiments, using crude extracts from cells grown under nitrogen-fixing conditions, indicated that an

upstream AT-rich region present in some *nif* promoters was protected from DNase I digestion (Beynon *et al.*, 1983). At the time, it was thought that this protection represented an interaction with RNA polymerase but later it was shown to be due to the binding of Integration Host Factor (IHF) (Hoover *et al.*, 1990) (see below).

Mutations located in invariant nucleotides in the –24 and –12 regions of *nif* promoters prevented transcriptional activation (Buck *et al.*, 1985; Khan *et al.*, 1986; Ow *et al.*, 1985). The spacing between the conserved GG and GC motifs in the –24 and –12 elements was found to be critical for promoter activity, reflecting a stringent spacing requirement (Buck, 1986). Some of the *nif* promoters, when cloned on multicopy plasmids, inhibited chromosomal *nif* expression in *K. pneumoniae*, resulting in a Nif$^-$ phenotype (Buchanan-Wollaston *et al.*, 1981b; Reidel *et al.*, 1983). This "multicopy effect" was attributed to titration of NifA by the excess promoters, thus preventing transcriptional activation of genomic *nif* promoters. Mutations that suppressed the multicopy effect were located either in the consensus –24 –12 region or in sequences further upstream (Brown and Ausubel, 1984; Buck *et al.*, 1985). Deletion analysis of the *nifH*, *nifU* and *nifB* promoters revealed that sequences upstream of –100 were required both for NifA-mediated activation and transcriptional activation. Sequence analysis of *nif* promoters identified an invariant TGT–N_{10}–ACA motif, which was proposed as a NifA binding site, designated as an Upstream Activator Sequence (UAS). Remarkably, the position and orientation of the UAS was not critical for promoter activity and it functioned when positioned up to a distance of 2 kb from the downstream –24 –12 element (Buck *et al.*, 1986).

At about the same time, it was found that activation of the *glnA* promoter by NtrC also occurred at a distance and that NtrC-binding sites could function far upstream of the promoter (Reitzer and Magasanik, 1986). NifA and NtrC were therefore brought into the limelight as eukaryotic-like transcriptional activators that bound to regulatory sequences similar to enhancers. This realization ultimately led to the classification of these transcriptional activators as Enhancer Binding Proteins (EBPs). Further analysis of the spatial requirements for UAS function showed that activation was face-of-the-helix dependent, indicating a stereospecific requirement for positioning of the activator with respect to the –24 –12 region. It was, therefore, proposed that NifA and NtrC activate transcription via a DNA looping mechanism (Buck *et al.*, 1987; Minchin *et al.*, 1989). However, activation of transcription was not entirely dependent on the UAS sequences because, for example, weak activation of the *nifH* promoter by NtrC had been detected (Buck *et al.*, 1985). In these cases, it was argued that neither activator binding at the UAS nor DNA loop formation are absolute prerequisites for transcriptional activation and that the activator might contact the downstream bound RNA polymerase from solution. Comparison of promoter sequences that were not strictly dependent on the UAS suggested that a run of T residues between –17 to –14 might be critical in the response of the promoter in the absence of the UAS. Conversion of this sequence in the *nifH* promoter from CCCT to TTTT suppressed the requirement for the UAS to be located on the same face of the helix with respect to the –24 –12 sequence and the promoter was far more responsive to a truncated form of NifA lacking the DNA-binding domain (Buck and Cannon, 1989). Whereas the binding of RNA

polymerase to the wild-type *nifH* promoter could not be detected by *in vivo* footprinting, protection was observed with the mutant promoter. These observations suggested that relatively weak binding of RNA polymerase to the downstream sequence coupled with stereospecific binding of the activator at the UAS, ensures the fidelity of activation, thus, ensuring that this *nif* promoter is specifically activated by NifA (Morett and Buck, 1989).

Nucleotide sequencing of the *K. pneumoniae* regulatory genes, *ntrA, ntrB, ntrC, nifL* and *nifA*, was completed between 1985 and 1987, but initially provided only limited clues to their function because these were the first genes of their class to be sequenced. Although NtrA was proposed to be a sigma factor and appeared to compete with σ^{70} when overexpressed *in vivo* (Merrick and Stewart, 1985), the sequence of *K. pneumoniae ntrA* showed that the encoded protein was not similar to other sigma factors. However, it did contain potential DNA-binding regions, which conceivably could be involved in promoter recognition (Merrick and Gibbins, 1985). Comparison of the NifA and NtrC sequences revealed a strongly conserved central domain and a C-terminal domain containing a helix-turn-helix motif proposed to be required for DNA binding (Buikema *et al.*, 1985; Drummond *et al.*, 1986). The role of the C-terminal domains of NtrC and NifA in DNA binding was established subsequently by mutagenesis and *in vivo* footprinting experiments (Contreras and Drummond, 1988; Morett and Buck, 1988). The homology between these two proteins suggested a common mechanism of transcriptional activation, commensurate with previous observations that, when overexpressed, NifA could substitute for NtrC at promoters normally activated by the latter (Drummond *et al.*, 1983; Merrick, 1983; Ow and Ausubel, 1983). However, the amino-terminal domain of NtrC was clearly different to that of NifA, but was homologous to diverse bacterial regulatory proteins, including OmpR from *E. coli* and Spo0A from *Bacillus subtilis* (Drummond *et al.*, 1986). The sequence of *ntrB* initially revealed no homologues (MacFarlane and Merrick, 1985), but the subsequent sequencing of *nifL* revealed a common C-terminal domain present in NtrB and other regulatory proteins, including EnvZ, PhoR, CpxA and CheA (Drummond and Wootton, 1987). These proteins all had a corresponding regulatory partner belonging to the NtrC family. The existence of these regulatory pairs of proteins had also been noted in Fred Ausubel's laboratory. It was proposed that they had evolved from a common ancestral system that transduced environmental signals from the C-terminal domain of one protein (*e.g.*, NtrB) to the N-terminal domain of its partner (*e.g.*, NtrC). These protein pairs were called two-component regulatory systems (Nixon *et al.*, 1986; Ronson *et al.*, 1987). In retrospect, it is fascinating that the discovery of two-component regulation arose from studies on nitrogen control of nitrogen fixation.

5.3. Biochemical Studies on Regulatory Proteins

5.3.1. Nitrogen Regulatory Proteins and the Discovery of σ^{54}

The E. coli NtrC protein was first purified in Boris Magasanik's laboratory in 1983 (Reitzer and Magasanik, 1983) and was subsequently isolated from Salmonella

typhimurium and K. pneumoniae (Ames and Nikaido, 1985; Hawkes et al., 1985). In all cases, the protein was shown to be a dimeric DNA-binding protein, which recognised sequences with dyad symmetry and the consensus sequence, 5'-TGCACTA$(N)_3$TGGTGCAA-3' (Ames and Nikaido, 1985; Dixon, 1984; Hawkes et al., 1985). The purified protein repressed transcription from both the ntrBC promoter and the upstream glnA promoter, which are both transcribed by σ^{70}-RNA polymerase holoenzyme (Hawkes et al., 1985; Reitzer and Magasanik, 1983). These results confirmed genetic studies, which indicated that the ntrC product acts as a repressor of transcription at these promoters, and provided an obvious mechanism to explain how NtrC mediates "negative" control, but gave no clues as to how this protein might activate transcription in concert with NtrA.

In late 1985, two exciting breakthroughs enabled this question to be addressed. Using a coupled *in vitro* transcription-translation system as an assay for activity, *S. typhimurium* NtrA was partially purified (Hirschman *et al.*, 1985). The purified NtrA-containing fraction activated transcription from the *glnA* promoter dependent on the addition of NtrC and *E. coli* core RNA polymerase. When σ^{70} was substituted for NtrA, no activation of the *glnA* promoter was detected. Conversely, NtrA could not substitute for σ^{70} at the *lac*UV5 promoter. These properties, and the observation that NtrA co-purified with RNA polymerase during the early stages of purification, strongly suggested that NtrA is an alternative RNA polymerase sigma factor (Hirschman *et al.*, 1985). At about the same time, the *E. coli* NtrA was purified and was demonstrated to bind core RNA polymerase. *In vitro* transcription experiments demonstrated that activation of the *glnA* promoter required NtrC and a mutant form of NtrB (NtrB2302), in addition to NtrA and core RNA polymerase (Hunt and Magasanik, 1985). These results suggested that NtrB is required to activate NtrC and also that the product of *ntrA* is a sigma factor. It was proposed that the name of *ntrA* should be changed to *rpoN* and its product designated as σ^{60} (Hunt and Magasanik, 1985). The latter was subsequently changed to σ^{54}, once the molecular weight had been more accurately derived from sequence data. Thus, an alternative sigma factor, which recognised the unique –24 and –12 consensus sequences present in *rpoN*-dependent promoters, had been discovered. This finding led to major new drives to understand the structure and function of this novel sigma factor (Merrick, 1993).

The purification of NtrB made it possible to investigate its role in modulating the activity of NtrC because previous genetic experiments had indicated that NtrB converts inactive NtrC into a form capable of activating transcription under nitrogen-limiting conditions. This response to nitrogen status involves a complex metabolic cascade mediated by the products of *glnB* (P_{II}) and *glnD* (uridylyl-transferase) (see Merrick, this volume). It was observed that, in a mixture containing NtrB, NtrC and ATP, NtrB catalyses the phosphorylation of NtrC and only the covalently modified form of the activator was competent to activate transcription at the *glnA*p2 promoter. When purified P_{II} protein was added to the mixture of NtrB and phosphorylated NtrC, NtrC became dephosphorylated and inactive as a transcriptional activator. These experiments suggested that NtrB is a protein kinase that phosphorylates NtrC to activate transcription and that the kinase

activity of NtrB is antagonised by P_{II} in response to nitrogen status (Ninfa and Magasanik, 1986). This experiment was the first to demonstrate phosphotransfer in a two-component regulatory system. Subsequently, it was shown that the amino-terminal domain of NtrC is phosphorylated by NtrB and that phosphorylated NtrC has an autophosphatase activity (Keener and Kustu, 1988).

Further mechanistic insights into the mechanism of action of the NtrC protein were achieved when it was observed that ATP hydrolysis was required for transcriptional activation. In the absence of NtrC, the σ^{54}-RNA polymerase holoenzyme was shown to form a stable complex at the *glnA* promoter, referred to as the closed complex (Popham *et al.*, 1989). NtrC was, therefore, not required to stabilise the binding of the σ^{54}-holoenzyme to the promoter but appeared to be required at a subsequent stage to catalyse the conversion of the closed complex to the open complex in which the DNA strands surrounding the transcription start site are melted prior to initiation. To simplify the analysis of transcriptional initiation, a mutant form of the NtrC protein, which was competent to activate transcription in the absence of NtrB, was used. This altered NtrC avoided the complication of the ATP requirement for phosphorylation of NtrC. However, when studied in the absence of NtrB, the mutant NtrC protein still required ATP to catalyse formation of open promoter complexes (Popham *et al.*, 1989). Because non-hydrolysable analogues did not substitute for ATP, it was concluded that ATP hydrolysis catalysed by NtrC is necessary for the isomerization of the closed complex between σ^{54}-holoenzyme and the *glnA* promoter to the open promoter complex. Further analysis of the mechanism of long distance activation involved whether the NtrC enhancer sites could function *in vitro* when located in *trans* on a separate plasmid to that carrying the σ^{54}-recognition sequence. Whereas the enhancer sites could function *in cis* when located on the same plasmid as the downstream promoter sequence, they only functioned *in trans* when present on different circles of a singly-linked catenane. This finding suggested that one function of the enhancer sequences is to "tether" the NtrC in the vicinity of the promoter and, therefore, increase the frequency with which it encounters σ^{54}-holoenzyme (Wedel and Kustu, 1991). DNA looped structures formed by interaction between enhancer-bound NtrC and σ^{54}-holoenzyme were observed directly in the electron microscope. The DNA loops were only observed when ATP was also present, implying that stable loops are only formed when polymerase has completed the transition to the open complex (Su *et al.*, 1990).

The rapid developments in understanding the function of the nitrogen regulatory proteins at the *glnA* promoter were soon utilised to probe the mechanism of activation of *nif* transcription by NtrC. *In vitro* experiments with the *K. pneumoniae* *nifLA* promoter demonstrated that NtrB, NtrC and σ^{54} are required for transcriptional activation and the upstream promoter region contains tandem binding sites for NtrC located at –142 and –163 (Austin *et al.*, 1987; Wong *et al.*, 1987). These sites do not show strong homology with the NtrC DNA-binding site consensus and were shown to have a low affinity for NtrC compared with sites in the *glnA* promoter (Minchin *et al.*, 1988). Consistent with this finding, *nifLA*