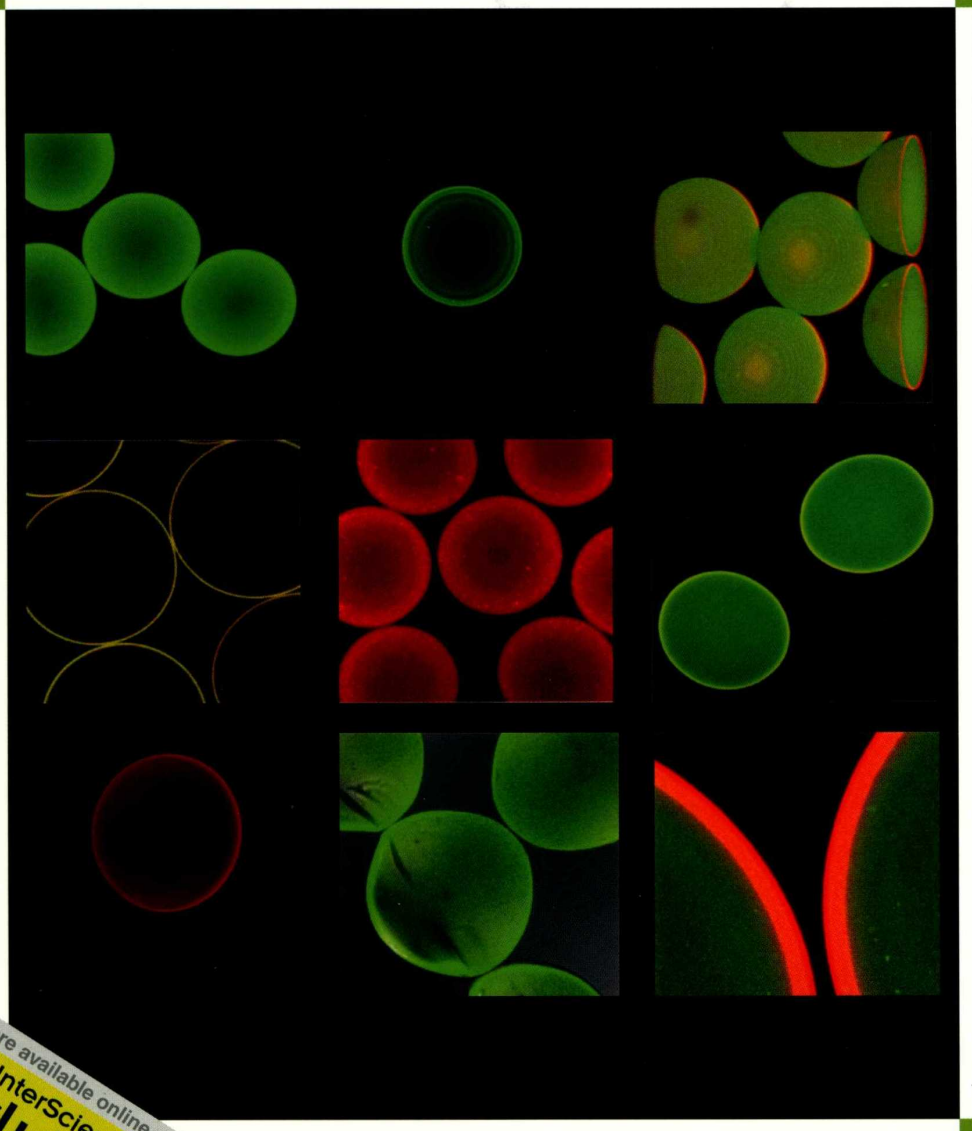
Ex. 1035 Trinh 2003

Volume 82, Number 4, May 20, 2003
Articles published online in Wiley InterScience 5 March 2003–6 March 2003

Biotechnology and Bioengineering



cles in this journal are available online ahead of the print issue
Wiley InterScience
EarlyView®
www.interscience.wiley.com

Biotechnology and Bioengineering

Editor
Douglas S. Clark
Department of Chemical Engineering, University of California Berkeley, California

Associate Editors

Michael J. Betenbaugh
Johns Hopkins University Baltimore, Maryland

James D. Bryers
University of Connecticut Farmington, Connecticut

Pauline Doran
University of New South Wales Sydney, Australia

Jonathan S. Dordick
Rensselaer Polytechnic Institute Troy, New York

Charles A. Haynes
University of British Columbia Vancouver, BC, Canada

Jeffrey A. Hubbell
ETH—Zürich and University of Zürich Zürich, Switzerland

Jay D. Keasling
University of California Berkeley, California

Lee R. Lynd
Dartmouth College Hanover, New Hampshire

Tadashi Matsunaga
Tokyo University of Agriculture and Technology Tokyo, Japan

Jens Nielsen
Technical University of Denmark Lyngby, Denmark

Advisory Board

Harvey W. Blanch
University of California Berkeley, California

Ho Nam Chang
Korea Advanced Institute of Science & Technology Taejon, Korea

George Georgiou
University of Texas Austin, Texas

Isao Karube
University of Tokyo Tokyo, Japan

Alexander M. Klibanov
Massachusetts Institute of Technology Cambridge, Massachusetts

Maria-Regina Kula
Heinrich-Heine Universität Düsseldorf in der KFA Jülich, Germany

Douglas A. Lauffenburger
Massachusetts Institute of Technology Cambridge, Massachusetts

Bernhard Ø. Palsson
University of California at San Diego La Jolla, California

Eleftherios T. Papoutsakis
Northwestern University Evanston, Illinois

Michael L. Shuler
Cornell University Ithaca, New York

John Villadsen
Technical University of Denmark Lyngby, Denmark

Daniel I.C. Wang
Massachusetts Institute of Technology Cambridge, Massachusetts

Editorial Board

Mohamed Al-Rubeai
Birmingham, United Kingdom

Kristi Anseth
Boulder, CO, USA

Jun-ichi Anzai
Sendai, Japan

Mark Arnold
Iowa City, IA, USA

William Bentley
College Park, MD, USA

Michael Butler
Winnipeg, MB, Canada

Manuel J.T. Carrondo
Oeiras, Portugal

Jeffrey Chalmers
Columbus, OH, USA

Howard Chase
Cambridge, United Kingdom

Julian Chaudhuri
Bath, United Kingdom

Wilfred Chen
Riverside, CA, USA

Bruce Dale
East Lansing, MI, USA

Scott Diamond
Philadelphia, PA, USA

Erik J. Fernandez
Charlottesville, VA, USA

Martin Fussenegger
Zurich, Switzerland

Peter Halling
Glasgow, United Kingdom

Hiroyuki Honda
Nagoya, Japan

Beth Junker
Rahway, NJ, USA

Michael Ladisch
West Lafayette, IN, USA

Sang Yup Lee
Daejeon, Korea

Abraham Lenhoff
Newark, DE, USA

Gunnar Lidén
Lund, Sweden

Robert Lortie
Montreal, Quebec, Canada

Brian McNeil
Glasgow, United Kingdom

William M. Miller
Evanston, IL, USA

Alvin Nienow
Birmingham, United Kingdom

Henrik Pedersen
Piscataway, NJ, USA

James Petersen
Pullman, WA, USA

James M. Piret
Vancouver, BC, Canada

Octavio T. Ramírez
Cuernavaca, Mexico

Govind Rao
Baltimore, MD, USA

Joseph O. Rich
Mt. Prospect, IL, USA

Anne Skaja Robinson
Newark, DE, USA

David Schaffer
Berkeley, CA, USA

Masayasu Suzuki
Toyama, Japan

Vytas Svedas
Moscow, Russia

Masahiro Takagi
Ishikawa, Japan

Owen Thomas
Lyngby, Denmark

Nigel Titchener-Hooker
London, United Kingdom

Luuk van der Wielen
Delft, The Netherlands

Mark van Loosdrecht
Delft, The Netherlands

William R. Wagner
Pittsburgh, PA, USA

Ping Wang
Akron, OH, USA

Mark Worden
East Lansing, MI, USA

Charles E. Wyman
Hanover, NH, USA

Jian-Jiang Zhong
Shanghai, China

Andrew Zydney
State College, PA, USA

Founding Editor: Elmer L. Gaden, Jr., *University of Virginia, Charlottesville, Virginia*

Biotechnology and Bioengineering (Print ISSN: 0006-3592; Online ISSN: 1097-0290 at Wiley InterScience, www.interscience.wiley.com) Volumes 81–84 are published 28 times a year, semi-monthly, except tri-monthly in March, June, September, and December, by Wiley Subscription Services, Inc., a Wiley Company, 111 River Street, Hoboken, NJ 07030.

Copyright © 2003 Wiley Periodicals, Inc., a Wiley Company. All rights reserved. No part of this publication may be reproduced in any form or by any means, except as permitted under section 107 or 108 of the 1976 United States Copyright Act, without written permission of the publisher, authorization through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400, fax 978-750-4470. Periodicals postage paid at Hoboken, NJ, and at additional mailing offices.

The copyright notice appearing at the bottom of the first page of an article in the journal indicates the copyright holder's consent that copies may be made for personal or internal use, or for the personal or internal use of specific clients, on the condition that the copier pay for copying beyond that permitted by Sections 107 and 108 of the United States Copyright Law. The per copy fee for each item is to be paid through the Copyright Clearance Center. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. Such permission requests and other permission inquiries should be addressed to the Permissions Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, http://www.wiley.com/go/permissions.

Subscription Price (Volumes 81–84, 2003): Print only: $4495.00 in U.S., $4775.00 in Canada and Mexico, $5013.00 outside North America. Electronic only: $4495.00 worldwide. A combination price of $4720.00 in U.S., $5000.00 in Canada and Mexico, $5238.00 outside North America, includes the subscription in both electronic and print formats. Personal rate: $595.00 in North America, and $763.00 rest of world. Subscriptions at the personal rate are available to individuals. All subscriptions containing a print element, shipped outside U.S., will be sent by air. Payment must be made in U.S. dollars drawn on U.S. bank. Claims for undelivered copies will be accepted only after the following issue has been delivered. Please enclose a copy of the mailing label. Missing copies will be supplied when losses have been sustained in transit and where reserve stock permits. Please allow four weeks for processing a change of address. For subscription inquiries, please call (201) 748-6645; E-mail: SUBINFO@wiley.com.

Postmaster: Send address changes to Biotechnology and Bioengineering. Subscription Distribution, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030.

Advertising Sales: Inquiries concerning advertising should be forwarded to Advertising Sales, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030; (201) 748-8832. Advertising Sales, European Contacts: Jackie Sibley, c/o John Wiley & Sons, Ltd., Baffins Lane, Chichester, West Sussex PO19 1UD, England. Tel: 44 1243 770 351; Fax: 44 1243 770 432; e-mail: adsales@wiley.co.uk.

Reprints: Reprint sales and inquiries should be directed to the Customer Service Department, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Tel: (201) 748-8789.

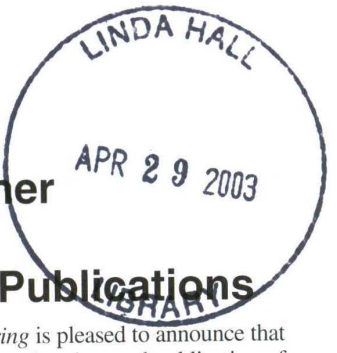
Manuscripts should be submitted to Biotechnology and Bioengineering, Editorial Receiving Office, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030. Phone: (201) 748-6297; Fax: (201) 748-8881; E-mail: biotechbioeng@wiley.com

All Other Correspondence: Address to Biotechnology and Bioengineering, Publisher, c/o John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030.

The contents of this journal are indexed or abstracted in Analytical Abstracts, BIOSIS Previews, Biological Abstracts, Biotechnology Citation Index®, Chemical Abstracts, Current Contents®/Agriculture, Biology, and Environmental Science, Current Contents®/Life Sciences, International Abstracts in Operations Research (IAOR), IAOR-Online, ISI Alerting Services (includes Research Alert®), Reference Update®, Science Citation Index®, and SciSearch® (also known as Science Citation Index-Expanded).

Editorial Production, John Wiley
Paul Dlugokencky

♾ **This paper meets the requirement of ANSI/NISO Z39.48-1992 (Permanence of Paper)**



Announcements from the Publisher

New: Online Submission and Peer Review

In taking a step toward facilitating seamless international scholarly communication, *Biotechnology and Bioengineering* is now pleased to offer web-based submissions and peer review.
To submit your manuscript online, please:

1. Prepare your manuscript and illustrations in appropriate format, according to the instructions given at http://bit-wiley.manuscriptcentral.com/ under the "Instructions and Forms" button. Please also be sure that your paper conforms to the scientific and style instructions of the journal, given below.
2. If you have not already done so, create an account for yourself in the system at the submission site, http://bit-wiley.manuscriptcentral.com/ by clicking on the "Create an Account" button. To monitor the progress of your manuscript throughout the review process, just login periodically and check your Author Center.
3. Please be sure to study the *Instructions and Forms* given at the site carefully, and then let the system guide you through the submission process. Online help is available to you at all times during the process. You are also to exit/re-enter the process at any stage before finally "submitting" your work. All submissions are kept strictly confidential. If you have any questions, do not hesitate to contact us at edsupport@wiley.com.

Editorial Receiving Office

The editorial receiving office for *Biotechnology and Bioengineering* has moved. Please direct all manuscript submissions, correspondence and inquiries to the following address:

Biotechnology and Bioengineering
Editorial Receiving Office
John Wiley & Sons, Inc.
111 River Street
Hoboken, NJ 07030

Phone: (201) 748-6297
Fax: (201) 748-8881
E-Mail: biotechbioeng@wiley.com

Accelerated Publications

Biotechnology and Bioengineering is pleased to announce that the journal will provide accelerated review and publication of *Accelerated Publications* reporting new findings of unusual and timely significance. Authors who wish their manuscript to be considered for accelerated publication should include a brief description of the reasons why the manuscript is suitable for "fast tracking" in their letter of transmittal accompanying the manuscript upon initial submission. The names, addresses (including e-mail), telephone, and fax numbers of five potential reviewers must also be included with the submission. The Editor and/or the Associate Editors will select manuscripts suitable for fast tracking. Fast tracked manuscripts will be reviewed within four weeks and published within three months from the date of the final acceptance.

Cover Illustrations

Biotechnology and Bioengineering would also like to encourage authors to submit color figures as candidates for cover illustrations when submitting manuscripts. These illustrations may be related to, but not necessarily included in, their manuscript. Alternatively, authors may submit figures which portray the fields of biotechnology and bioengineering at the molecular, cellular, bench-top, or production level as possible cover illustrations in keeping with the general themes of the journal. Authors should submit a brief description of their illustration if it is not included in their manuscript.

Abstracting and Indexing

All articles published in *Biotechnology and Bioengineering* will be included in MedLine and PubMed. The Journal is also covered by the following indexing and abstracting services:
Analytical Abstracts
BIOSIS Previews, Biological Abstracts
Biotechnology Citation Index®
Chemical Abstracts
Current Contents®/Agriculture, Biology, and Environmental Science
Current Contents®/Life Sciences
International Abstracts in Operations Research (IAOR)
IAOR-Online
ISI Alerting Services (includes Research Alert®)
Reference Update®
Science Citation Index®
SciSearch® (also known as Science Citation Index-Expanded).

The Gaden Award

Biotechnology and Bioengineering and John Wiley & Sons, Inc. is pleased to announce that, in honor of its founding editor Elmer L. Gaden, Jr., the journal provides an annual award for a high-impact paper reflecting exceptional innovation, creativity, and originality. Each year the winner of this award receives an honorarium and travel stipend to attend the American Chemical Society national meeting where he or she is invited to speak at an ACS Division of Biochemical Technology session. To nominate a paper for this prestigious award, please contact Dr. Douglas S. Clark, editor.

2001 Gaden Award

The 2001 Gaden Award was presented to Professor Govind Rao for:
Low-cost microbioreactor for high throughput bioprocessing
Yordan Kostov, Peter Harms, Lisa Randers-Eichhorn, Govind Rao
Biotechnology and Bioengineering 72:346–352, 2001

2000 Gaden Award

The 2000 Gaden Award was presented to Professor James E. Bailey for:
Engineering of coordinated up- and down-regulation of two glycosyltransferases of the o-glycosylation pathway in Chinese Hamster Ovary (CHO) cells
Elisabetta G.P. Prati, Mattia Matasci, Tobias B. Suter, Andre Dinter, Andriana R. Sburlati, James E. Bailey
Biotechnology and Bioengineering 68:239–244, 2001

1999 Gaden Award

The 1999 Gaden Award was presented to Professor Sang Yup Lee for:
Chiral compounds from bacterial polyesters: Sugars to plastics to fine chemicals
San Yup Lee, Young Lee, Fulai Wang
Biotechnology and Bioengineering 65:363–368, 1999

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Biotechnology and Bioengineering

Volume 82, Number 4, May 20, 2003

Contents

379 Simple and Efficient Solid Support Scavenging of Excess Acyl Donors After Enzymatic Acylations in Organic Solvents
A.Y. Usyatinsky, N.M. Astakhova, and Y.L. Khmelnitsky
Published online 5 March 2003

386 Visualization of Alginate–Poly-L-lysine–Alginate Microcapsules by Confocal Laser Scanning Microscopy
B.L. Strand, Y.A. Mørch, T. Espevik, and G. Skjåk-Bræk
Published online 5 March 2003

395 Determination of Laccase Activity in Mixed Solvents: Comparison Between Two Chromogens in a Spectrophotometric Assay
G. Cantarella, F. d'Acunzo, and C. Galli
Published online 5 March 2003

399 Laser Impact Assessment in a Biofilm-Forming Bacterium *Pseudoalteromonas carrageenovora* Using a Flow Cytometric System
K. Nandakumar, H. Obika, T. Shinozaki, T. Ooie, A. Utsumi, and T. Yano
Published online 5 March 2003

403 High-Density Seeding of Myocyte Cells for Cardiac Tissue Engineering
M. Radisic, M. Euloth, L. Yang, R. Langer, L.E. Freed, and G. Vunjak-Novakovic
Published online 5 March 2003

415 Halogenated Monoterpene Production by Microplantlets of the Marine Red Alga *Ochtodes secundiramea* Within an Airlift Photobioreactor Under Nutrient Medium Perfusion
J.P. Polzin and G.L. Rorrer
Published online 5 March 2003

429 Microfiltration and Ultrafiltration of Polysaccharides Produced by Fermentation Using a Rotating Disk Dynamic Filtration System
A. Brou, M.Y. Jaffrin, L.H. Ding, and J. Courtois
Published online 5 March 2003

438 Effect of Methanol Feeding Strategies on Production and Yield of Recombinant Mouse Endostatin From *Pichia pastoris*
L.B. Trinh, J.N. Phue, and J. Shiloach
Published online 5 March 2003

(continued on the next page)

This Journal is Online
WILEY InterScience®
www.interscience.wiley.com

Volume 82, Number 4 was mailed the week of March 31, 2003.

(continued from the previous page)

445 Glucose-6-Phosphate Dehydrogenase Partitioning in Two-Phase Aqueous Mixed (Nonionic/Cationic) Micellar Systems
C.O. Rangel-Yagui, H. Lam, D.T. Kamei, D.I.C. Wang, A. Pessoa, Jr., and D. Blankschtein
Published online 5 March 2003

457 Role of Pericellular Matrix in Development of a Mechanically Functional Neocartilage
R.D. Graff, S.S. Kelley, and G.M. Lee
Published online 5 March 2003

465 Electrochemical Optical Waveguide Lightmode Spectroscopy (EC-OWLS): A Pilot Study Using Evanescent-Field Optical Sensing Under Voltage Control to Monitor Polycationic Polymer Adsorption Onto Indium Tin Oxide (ITO)-Coated Waveguide Chips
J.P. Bearinger, J. Vörös, J.A. Hubbell, and M. Textor
Published online 6 March 2003

474 Levitation and Movement of Human Tumor Cells Using a Printed Circuit Board Device Based on Software-Controlled Dielectrophoresis
L. Altomare, M. Borgatti, G. Medoro, N. Manaresi, M. Tartagni, R. Guerrieri, and R. Gambari
Published online 6 March 2003

480 Fusion Protein of *Vitreoscilla* Hemoglobin With D-Amino Acid Oxidase Enhances Activity and Stability of Biocatalyst in the Bioconversion Process of Cephalosporin C
Y.-H. Khang, I.-W. Kim, Y.-R. Hah, J.-H. Hwangbo, and K.-K. Kang
Published online 6 March 2003

489 Monitoring of a Sequencing Batch Reactor Using Adaptive Multiblock Principal Component Analysis
D.S. Lee and P.A. Vanrolleghem
Published online 6 March 2003

498 Forthcoming Articles

Effect of Methanol Feeding Strategies on Production and Yield of Recombinant Mouse Endostatin From *Pichia pastoris*

L. B. Trinh, J. N. Phue, Joseph Shiloach

Biotechnology Unit, National Institutes of Health, NIDDK Building 6, Room B1-33, Bethesda, Maryland 20892-2715; telephone: 301-496-9719; fax: 301-451-5911; e-mail: yossi@nih.gov

Received 21 April 2002; accepted 14 October 2002

Published online 5 March 2003 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/bit.10587

Abstract: *Pichia pastoris,* a methylotrophic yeast, is an efficient producer of recombinant proteins in which the heterologous gene is under the control of the methanol-induced *AOX1* promoter. Hence, the accepted production procedure has two phases: In the first phase, the yeast utilizes glycerol and biomass is accumulated; in the second phase, the yeast utilizes methanol which is used both as an inducer for the expression of the recombinant protein and as a carbon source. Since the yeast is sensitive to methanol concentration, the methanol is supplied gradually to the growing culture. Three methanol addition strategies were evaluated for the purpose of optimizing recombinant endostatin production. Two strategies were based on the yeast metabolism; one responding to the methanol consumption using a methanol sensor, and the other responding to the oxygen consumption. In these two strategies, the methanol supply is unlimited. The third strategy was based on a predetermined exponential feeding rate, controling the growth rate at 0.02 h^{-1}, in this strategy the methanol supply is limited. Throughout the induction phase glycerol, in addition to methanol, was continuously added at a rate of 1 g L h^{-1}. Total endostatin production was similar in all three strategies, (400 mg was obtained from 3 L initial volume), but the amount of methanol added and the biomass produced were lower in the predetermined rate method. This caused the specific production of endostatin per biomass and per methanol to be 2 times higher in the predetermined rate than in the other two methods, making the growth control strategy not only more efficient but also more convenient for downstream processing. © 2003 Wiley Periodicals, Inc. **Biotechnol Bioeng* **82**: 438–444, 2003.
Keywords: *Pichia pastoris;* growth strategies; methanol; endostatin

INTRODUCTION

Pichia pastoris is an efficient producer of recombinant proteins (Cregg et al., 2000). This microorganism is a methylotrophic yeast (Faber et al., 1995) that utilizes methanol as its carbon source and grows to very high cell density (above 50% wet weight concentration). In many cases it secretes the produced protein to the culture media. The reason for the efficiency of this yeast as a producer of recombinant protein is the biosynthesis of alcohol oxidase (AOX). AOX is the first enzyme in the methanol oxidation pathway. The enzyme is produced in response to the methanol presence in the growth media. Methanol initiates the enzyme biosynthesis by activating the *AOX1* promoter, causing the enzyme accumulation to about 30% of the total yeast proteins. Therefore, the gene of choice is introduced downstream to the alcohol oxidase promoter and is expected to express the desired protein in an equivalent amount (Cereghino and Cregg, 2000).

Three types of recombinant Pichia can be obtained (Higgins and Cregg, 1998): Mut^+ where the two methanol oxidases genes *AOX1* and *AOX2* are intact; Mut^S, where only *AOX2*, which is responsible for 15% of the protein biosynthesis, is intact; and Mut^- where both *AOX1* and *AOX2* genes are disrupted. The Mut^+ strain is the most responsive to methanol concentration and is the recombinant strain most commonly used.

Alcohol oxidase promoter is activated by methanol, but it is also repressed by carbon source such as glycerol or glucose (Tschopp et al., 1987). In addition, the microorganism is sensitive to methanol, and concentrations above 0.4% inhibit growth and production (Katakuga et al., 1998; Zhang et al., 2000a). Hence, the accepted production procedure from the Mut^+ strain has two phases. In the first phase, the yeast utilizes glycerol and biomass is accumulated. In the second phase, the glycerol supply is terminated or reduced (Brierley et al., 1990; d'Anjou and Daugulis, 1997), methanol alone or in combination with lower concentration of glycerol (mixed feed) is added and the biosynthesis of the recombinant protein is initiated.

Since the yeast is sensitive to methanol concentration, methanol needs to be supplied continuously to the growing culture keeping its concentration below 0.4%. Several fed-batch strategies for methanol addition have been established (Zhang et al., 2000b). These strategies can be metabolism related, based on parameters such as methanol consumption (Curvers et al., 2001; Guarna et al., 1997; Hellwig et al.,

Correspondence to: Joseph Shiloach

© 2003 Wiley Periodicals, Inc. * This article is a US Government Work and, as such, is in the public domain in the United States of America.

This material may be protected by Copyright law (Title 17 U.S. Code)

2000; Katakuga et al., 1998; Kobayashi et al., 2000), oxygen consumption (Byrne et al., 2000; Chung, 2000; Minning et al., 2001), pH control or CO_2 concentration. While implementing these strategies, the microorganism itself is controling the amount of methanol added as a response to its own metabolism, the methanol supply is practically unlimited and its concentration can have any value, usually not more than 4 g/L^{-1}. Growth strategies can also be based on predetermined methanol addition schemes at constant, linear, or exponential rates (Chauhan et al., 1999; Freyre et al., 2000; Inan et al., 1999; Marasugi et al., 2000, Zhang et al., 2000a). In these strategies the methanol feeding rate and not the microorganism is controling the growth, the actual methanol concentration is zero and its supply is limited.

It is very likely that the growth and production characteristics of the culture such as growth rate, production yield, and methanol consumption would depend on the fed-batch strategy. It is therefore important to evaluate the various fed-batch methods when the production process of a recombinant protein is being evaluated. In some cases, metabolism-related method is preferred (Curvers et al., 2001), and in other cases, nonmetabolism-related method, in which the growth rate can be controlled, is preferred (Zhang et al., 2000a). However, detailed comparisons of various fed-batch strategies for the production of a specific recombinant protein are very limited (Kupcsulik et al., 2001; Minning et al., 2001; Zhang et al., 2000b). Our purpose here is to evaluate several growth strategies for the production of extracellular murine endostatin. Endostatin, is a 20 KDa fragment of collagen XVIII, that had been shown to have an inhibitory effect on angionenesis and therefore can potentially be used as a tumor growth suppressor (Figg et al., 2002; O'Reilly et al., 1997). The compound is currently being tested in humans and different delivery strategies are being evaluated in mice. Gram quantities of the compound were needed and therefore efficient production procedure was required.

The growth strategies evaluation performed in this work can provide information not only on growth and production, but also on the specific production values of biomass in relation to methanol, and the specific production values of protein in relation to biomass and methanol.

MATERIAL AND METHODS

Yeast Strain and Fermentation Processes

Pichia pastoris strain GS 115 His^+ Mut^+ expressing mouse endostatin was prepared as described earlier (Boehm et al., 1999).

Bench-top fermentation was performed in a 7-L fermentor (New Brunswick Scientific, Edison, NJ) The fermentor is interfaced to an MD-Biostat System (B. Braun Biotech USA, Allentown, PA) furnished with an adaptive control algorithm (Hsiao et al., 1992) that maintained dissolved oxygen at 30% saturation by adjusting the agitation and air or oxygen. The fermentor was also equipped with a methanol sensor (Trinh et al., 2000; Wagner et al., 1997). Three liters BMGY media (peptone, 2%; yeast extract, 1%; YNB, 1.34%; glycerol, 3%; biotin, 4×10^{-5}%; and 100 m*M* potassium phosphate pH 6.0) was inoculated with 100 mL of overnight culture (O.D. $\approx$ 20 at 600 nm). The culture was grown in batch mode for approximately 16 h (O.D. $\approx$ 60 at 600 nm), after which fed-batch mode was started by adding a 50% glycerol solution at a flow rate of 18 mL L^{-1} h^{-1}; during this period, a dose of 0.3% (w/v) methanol was also pumped into the fermentor to calibrate the methanol sensor. After stabilization of the sensor signal (around 1 h), the glycerol feed rate was reduced gradually over a 1-hour period to 2 mL L^{-1} h^{-1}, allowing the culture to adapt smoothly to methanol. The reduced glycerol feed rate was kept constant throughout the methanol induction period (mixed feed). After the adaptation period, methanol was added according to the selected fed-batch strategy. The duration of the methanol addition phase was 48 h. During the fermentation the pH was kept at 6.5 using 50% NH_4OH solution, and foam was controlled by the addition of antifoam 289 (Sigma Chemical Co., St. Louis, MO.)—this antifoam did not affect growth or endostatin production.

Pilot-scale fermentation was performed in 80-L fermentor (Bioflo 5000, New Brunswick Scientific, Edison, NJ) equipped with the same control and data acquisition as the bench top.

Methanol Addition Strategies

Fed-Batch Fermentation Based on On-Line Methanol Sensor

After the adaptation period, the addition of methanol containing 12 mL/L^{-1} PTMI solution (per L: cupric sulphate-$5H_2O$ 6.0 gm; sodium iodide 0.08 gm; manganese sulphate-H_2O 3.0 gm; sodium molibdate-$2H_2O$ 0.2 gm; boric acid 0.02 gm; cobalt chloride 0.5 gm; zinc chloride 20.0 gm; ferrous sulphate-$7H_2O$ 65.0 gm; biotin 0.2 gm; sulfuric acid 5.0 gm) was initiated. Methanol concentration was controlled at 3.0 g/L^{-1} by an on-line monitoring and control device described previously (Trinh et al., 2000; Wagner et al., 1997).

Fed-Batch Fermentation Based on Dissolved Oxygen Signal

After the adaptation period, methanol solution containing 12 mL/L^{-1} PTMI was added into the fermentor when the dissolved oxygen signal spiked 10% above the set-point. Each time methanol was added to adjust the methanol concentration to 0.3%.

Fed-Batch Fermentation Based on Predetermined Exponential Feed Rate

After the adaptation period, methanol was pumped to the culture by a ChemTec pump (Scilog, Middleton, WI) at a

rate based on the following relationship (Zhang et al., 2000a), $F = 0.962\ \mu X_0 V_0 e^{\mu t}$ where F is the methanol flow rate (g/L^{-1}), X_0 is the biomass-wet weight at induction ($gWCW/L^{-1}$), V_0 is culture volume at induction (L), μ is culture growth rate under methanol (h^{-1}).

The methanol flow rate was ramped exponentially to keep the culture growth rate at 0.02 h^{-1}.

Analytical Methods

Off-Line Measurements

Off-line methanol measurements were made using YSI Biochemistry Analyzer (Yellow Spring Instruments, Yellow Spring, OH).

Endostatin Determination

Samples taken during the course of induction were centrifuged at 4000*g* for 40 min. The supernatant was diluted to adjust the conductivity to 5.5 mS/cm and the pH to 6.0 and was loaded on a 5 mL HiTrap™ SP HP (Amersham Pharmacia Biotech AB, Piscataway, NJ) at a ratio of 20 mg total protein per mL packed resin. The column was washed with 10 column volumes of 20 m*M* phosphate buffer pH 6.0 containing 50 m*M* NaCl buffer, and the endostatin was eluted with 3 column volumes of 20 m*M* phosphate containing 1.5*M* NaCl. Endostatin was analyzed by loading the various samples and increasing amount (0–8.6 μg) of known endostatin standard on 4–20% Tris-glycine SDS PAGE under reducing conditions and running at 125 volts for 90 min. After staining with Simply Blue™ SafeStain (Invitrogen, Carlsbad, CA) the amount of endostatin was determined by digitising and calculating with NIH image analysis software (NIH, Bethesda, MD).

Dry Weight Determination

Dry weight determination was done by drying cell paste in a moisture determination balance (Ohaus MB 200, NJ) for 3 h at 95°C.

Statistical Analysis

Each of the methanol feeding strategies was tested in three repetitions. Standard deviation was calculated for the average values of the three experiments from each feeding strategy.

RESULTS

Effect of the Various Fed-Batch Strategies on Growth Rate, Biomass Accumulation, and Methanol Supply

Growth rate and biomass accumulation during the three different fed-batch strategies are shown in Figure 1 and Table I. The growth rate during the predetermined methanol addition strategy was kept constant at a value of 0.02 h^{-1}, but the growth rates during the DO control strategy and the methanol sensor control strategy followed different patterns. The rate was 0.06 h^{-1} and 0.05 h^{-1}, respectively 10 h after the induction, and the rate continuous to decline throughout the induction phase which lasted 48 h. The biomass accumulation rate was corresponded to the growth rate. It accumulated at a constant rate during the predetermined addition strategy reaching a final level of 434 g DCW (dry cell weight). During the DO control strategy and the methanol sensor strategy, the biomass accumulation was much faster at the beginning of the induction phase, and slowed down towards the end of the induction period, reaching a value of 986 and 988 gDCW, respectively, which is over 2 times as much as accumulated in the predetermined growth rate strategy.

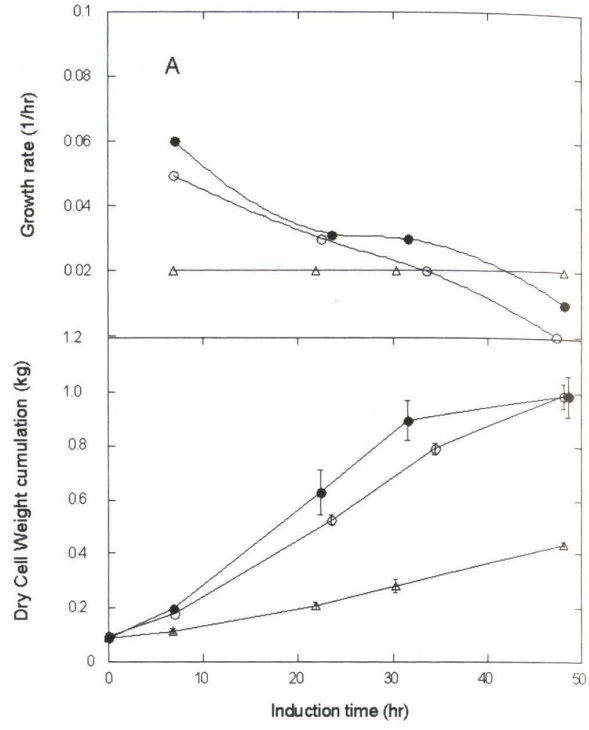


Figure 1. *Pichia pastoris* growth in response to different methanol supplies strategies. On line sensor (○), DO signal (●), predetermined rate (△). (Time zero indicates the initiation of the methanol induction phase). (A) Growth rate, (B) Biomass accumulation. Fermentation was conducted in a 7-L fermentor with initial volume of 3.0 L. Methanol addition was started when the biomass concentration was between 100 and 120 gram per liter (wet weight).

Methanol addition is described in Figure 2 and Table I. When the methanol concentration in the media was controlled by an online sensor, the total amount added was 3.23 L, compared with 2.25 L added in response to the dissolved oxygen concentration and to less than 1 L when the methanol was added using the predetermined rate.

Table I. Comparison of the various methanol feeding strategies. Fermentations were conducted in a bench top fermentor with initial volume of 3 L, the numbers reflect the values after 48 h induction.

Methanol control strategy	Final culture volume (L) V_0 = 3.0 L	Endostatin (mg/L culture)	Endostatin Total (mg)	Biomass (gDCW/L)	Biomass Total g DCW	Methanol consumed (L)	Base consumed (L)
Fed-batch using D.O. signal	6.09	66 ± 4	402	162 ± 2.6	986	2.25 ± 0.04	0.84
Fed-Batch using online gas sensor	7.06	67 ± 19	470	140 ± 2.7	988	3.23 ± 0.15	0.83
Fed-batch using predetermined feed rate	4.25	96 ± 16	408	102 ± 1.7	434	0.90 ± 0.11	0.34

Effect of the Various Fed-Batch Strategies on Endostatin Production

Total and volumetric endostatin productions during the different fed batch strategies are described in Figure 3 and Table I. The total amount of endostatin produced was similar in the fed-batch strategies using the dissolved oxygen signal and the predetermined feed rate. Approximately 400 mg were accumulated using these two strategies. Higher endostatin, 470 mg, was accumulated using the on-line sensor fed-batch strategy. Endostatin concentration was higher in the predetermined methanol feeding strategy compared with the dissolved oxygen control and the methanol sensor strategies. Final concentration of 96 mg/L was detected with the predetermined feed rate, compare with concentrations of 66 mg/L and 67 mg/L in the other two strategies.

Specific Production Values, Specific Consumption Values, and Production Rates in the Three Different Growth Strategies

Specific production of endostatin based on methanol utilization, biomass production based on carbon utilization (methanol and glycerol), and specific production of endostatin based on biomass are summarized in Table II. Specific endostatin production on methanol with the predetermined feeding rate was much higher than it was with the two other fed-batch strategies. It was 489 mg/L of methanol with the predetermined feeding rate compared with 160 mg/L, and 116 mg/L with the dissolved oxygen controlled strategy and the methanol sensor strategy, respectively. 0.72 mg endostatin per gram dry cell weight were produced with the predetermined feeding rate strategy compared with 0.31 and 0.33 mg endostatin per gram dry cell weight with the two other feeding strategies. The specific biomass production was 1.67 gram dry cell weight per gram carbon used (methanol and glycerol) with the predetermined feeding rate and 1.34 and 0.96 gram dry cell weight per gram carbon used, at the dissolved oxygen-controlled and the sensor-controlled strategies, respectively.

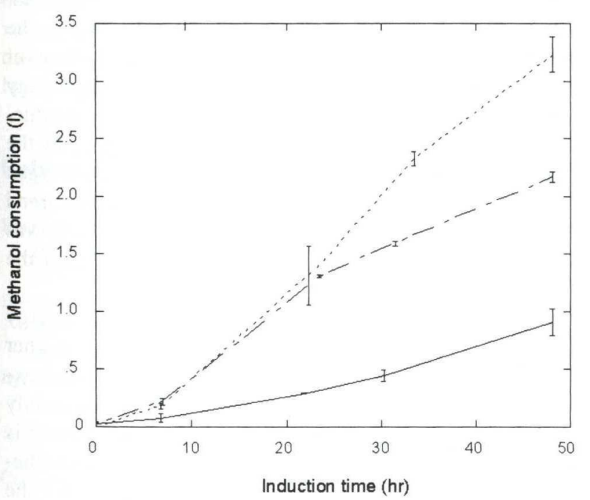


Figure 2. Methanol consumption in response to different supply strategies. On line sensor (– – – –), DO signal (— - — –), predetermined rate (—). (Time zero indicates the initiation of the methanol induction phase).

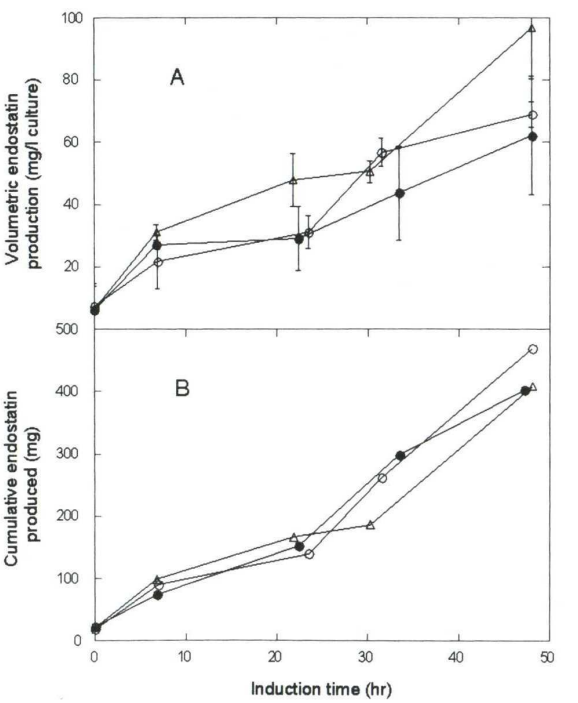


Figure 3. Endostatin production in response to different methanol supplies strategies. On-line sensor (○), DO signal (●), predetermined rate (△). (Time zero indicates the initiation of the methanol induction phase). (A) Total endostatin production, (B) volumetric endostatin production.

Table II. Specific endostatin and biomass production at different methanol feeding strategies. The yield constants reflect average values during the last 24 h of the induction phase.

Methanol control strategy	Specific endostatin production (mg Endostatin/ l methanol)	Specific endostatin production (mg Endostatin/ g DCW)	Specific biomass production (g DCW/g carbon)[a]
Fed-batch using D.O. signal	160 ± 6	0.31 ± 0.03	1.34 ± 0.11
Fed-batch using on-line gas sensor	116 ± 15	0.33 ± 0.09	0.96 ± 0.07
Fed-batch using predetermined feed rate	489 ± 8	0.72 ± 0.01	1.67 ± 0.16

[a]Specific biomass production was based on the total utilization of carbon in the methanol and glycerol.

Pilot production of endostatin using predetermined feeding rate

Based on the above information, a pilot-scale production (10 times bigger) was conducted using the preferred growth conditions. Typical process is shown in Figure 4; after 48 h induction 6.8 L of methanol were added, producing 2.7 g endostatin and 9.3 kg biomass (wet weight). Specific endostatin production yield was 25% lower than the value obtained in the bench-top fermentor, total consumption of methanol and total biomass production were also 25% lower.

DISCUSSION

Due to the sensitivity of *P. pastoris* to methanol concentration, production of recombinant proteins under the control of the *AOX1* promoter is done by implementing different fed-batch strategies controling the methanol concentration. Growth characteristics can be different at the various strategies, and therefore, recombinant protein production can also be affected. In addition, different values of metabolic indicators (e.g., methanol concentration, dissolved oxygen concentration) and different methanol feeding rates and profiles can be selected. These parameters can potentially affect growth and recombinant protein production within each strategy. Establishing optimal protein expression protocols should therefore include the evaluation of various growth and induction strategies in addition to other parameters such as pH and temperature. In this work, we examine the effect of three fed-batch strategies on the production of recombinant murine endostatin, evaluating not only the amount of endostatin produced but also the methanol consumption, biomass production, and the specific production values. Two strategies were based on the metabolic activity of the yeast (methanol and oxygen consumption) and one was based on supplying methanol at a predetermined exponential rate (limited methanol supply). The total production of endostatin with the three growth strategies was similar. Approximately 400 mg of endostatin were produced from a 3 L initial volume, but the amount of methanol used and the biomass produced were different, affecting final volume and specific production values of both endostatin and biomass.

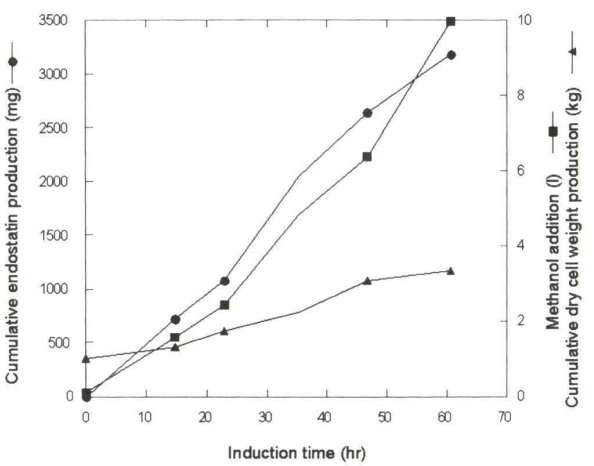


Figure 4. Pilot scale production of endostatin using predetermined methanol addition strategy: Endostatin production (●), methanol addition (■), biomass production (▲). (Time zero indicates the initiation of the methanol induction phase).

Feeding the methanol at a predetermined exponential profile and controling the growth rate at 0.02 h^{-1} was more efficient than the other two strategies. The amount of methanol used and the biomass produced were lower, causing the specific production of endostatin per biomass unit to be over 2 times higher than the dissolved oxygen controlled strategy or the sensor-based controlled strategy. Endostatin production per methanol used was almost 3 times higher in the predetermined strategy compared with the dissolved oxygen controlled strategy or the sensor strategy. A similar trend was observed with the specific biomass production. It was higher with the predetermined rate strategy than with the other two.

Methanol, the main carbon source is used both for energy generation and anabolism (Jahic et al., 2002). The higher efficiency of methanol utilization at the predetermined exponential rate suggests that the methanol is directed mainly towards energy generation, and that only a small portion is directed to biomass production. Since endostain biosynthesis is controlled by the same promoter responsible for the synthesis of the methanol oxidase, the amount of endostatin produced should be related to the methanol oxidase produced and should not be affected by the low methanol uti-

lization. It was also mentioned (Cregg, 1999) that when cells are fed methanol at a growth-limiting rate, the *AOX1* is induced to levels 3 to 5 times higher than in cells growing in excess methanol. Methanol consumption and biomass production were lower in the dissolved oxygen-controlled strategy than in the sensor-controlled strategy. This fact supports the above explanation since the methanol supply is not continuously unlimited as it is in the sensor controlled strategy.

Throughout the induction phase glycerol was supplied continuously to the culture at the rate of 1 gram per liter per hour. At this rate there was no glycerol accumulation. This supplemental carbon source is channelled directly to anabolism, (Jahic et al., 2002) and therefore compensate for possible shortage in carbon source as a result of methanol utilization for energy. While the glycerol supply may not be essential for the culture grown at excess methanol, it may be essential for the culture grown at limited methanol supply.

The results obtained from this comparison indicate that a growth strategy based on predetermined exponential feeding of methanol has an advantage for the production process of recombinant endostatin. The amount of methanol used and the biomass accumulation are lower than the values obtained using the dissolved oxygen or the methanol sensor methods, simplifying the endostatin recovery. In addition, the controlled growth rate limited the oxygen consumption and the heat production. It seems that these advantages are not necessarily specific for recombinant endostatin production. It is expected that any recombinant protein under the control of the *AOX1* promoter should respond in the same manner.

The high biomass concentration, above 50% wet weight, produced during the *Pichia pastoris* fermentation interferes with the recovery of the recombinant protein produced (Thommes et al., 2001; Trinh et al., 2000), especially when the protein is secreted to the media. In this aspect the predetermined addition methanol strategy offers another advantage, in addition to the high productivity, it simplifies the recovery process of the secreted endostatin due to the lower biomass produced and the lower overall volume.

References

Boehm T, Pirie-Shepherd S, Trinh LB, Shiloach J, Folkman J. 1999. Disruption of the KEX1 gene in *Pichia pastoris* allows expression of full-length murine and human endostatin. Yeast 15:563–572.

Brierley RA, Siegel RS, Bussineau CM, Craig WS, Holtz GC, Davis GR, Buckholtz RG. 1990. Mixed feed recombinant yeast fermentation. Patent Cooperation Treaty No. WO 90/03431 US.

Byrne MP, Titball RW, Holley J, Smith LA. 2000. Fermentation, purification, and efficacy of a recombinant vaccine candidate against botulinum neurotoxin type F from *Pichia pastoris*. Prot Expr Puri 18: 327–337.

Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast Pichia pastoris. FEMS Microbiol Rev 24:45–66.

Chauhan AK, Arora D, Khanna N. 1999. A novel feeding strategy for enhanced protein production by fed-batch fermentation in recombinant *Pichia pastoris*. Proc Biochem 34:139–145.

Chung JD. 2000. Design of metabolic feed controllers: Application to high density fermentation of *Pichia pastoris:* Mathematical model for fed-batch high cell density culture. Biotechnol Bioeng 68:298–307.

Cregg JM, Cereghino JL, Shi J, Higgins DR. 2000. Recombinant protein expression in *Pichia pastoris*. Molec Biotechnol—Part B 16:23–52.

Cregg JM. 1999. Expression in the methylotrophic yeast *Pichia pastoris*. In: Fernander JM, Hoeffler JP, editor. Gene expression systems. San Diego: Academic Press. p. 157–191.

Curvers S, Brixius P, Klauser T, Thommes J, Weuster-Botz D, Takors R, Wandrey C. 2001. Human chymotrypsinogen B production with *Pichia pastoris* by integrated development of fermentation and downstream processing. Part 1. Fermentation. Biotechnol Prog 17:495–502.

d'Anjou MC, Daugulis AJ 1997. A model-based feeding strategy for fed-batch fermentation of recombinant *Pichia pastoris*. Biotech Tech 11: 865–868

Faber KN, Harder W, AB G, Veenhuis M. 1995. Review: Methylotrophic yeasts as factories for the production of foreign proteins. Yeast 11: 1331–1344.

Figg WD, Kruger, EA, Price DK, Kim S, Dahunt WD. 2002. Inhibition of angiogenesis: Treatment potions for patients with metastatic prostate cancer. Inv New Drugs 20:183–194.

Freyre FM, Vazquez JE, Ayala M, Canaan-Haden L, Bell H, Rodriguez I, Gonzalez A, Cintado A, Gavilondo JV. 2000. Very high expression of an anti-carcinoembryonic antigen single chain Fv antibody fragment in the yeast *Pichia pastoris*. J Biotechnol 76:157–163.

Guarna MM, Lesnicki GJ, Tam BM, Robinson J, Radziminski CJ, Hasenwinkle D, Boraston A, Jervis E, Macgillivray RTA, Turner RFB, Kilburn DG. 1997. On-line monitoring and control of methanol concentration in shake-flask cultures of *Pichia pastoris*. Biotechnol Bioeng 56:297–286.

Hellwig S, Emde F, Raven NPG, Henke M, Van der Logt P, Fischer R. 2000. Analysis of single chain antibody production in *Pichia pastoris* using on-line methanol control in fed-batch and mixed-feed fermentation. Biotechnol Bioeng 74:344–352.

Higgins DR, Cregg JM. 1998. Pichia protocols. Methods in molecular biology. Totowa, NJ: Humana Press Inc.

Hsiao J, Ahluwalia M, Kaufman JB, Clem TR, Shiloach J. 1992. Adaptive control strategy for maintaining dissolved oxygen concentration in high density growth of recombinant *E. coli*. Ann NY Acad Sci 665: 320–333.

Inan M, Chiruvolu V, Eskridge KM, Vlasuk GP, Dickerson K, Brown S, Meagher MM. 1999. Optimization of temperature-glycerol-pH conditions for a fed-batch fermentation process for recombinant hookworm (*Ancylostoma caninum*) anticoagulant peptide (AcAP-5) production by *Pichia pastoris*. Enzyme MicrobTechnol 24:438–45.

Jahic M, Rotticci-Mulder JC, Matinelle M, Hult K, Enfors S-O. 2002. Modeling of growth and energy meyabolism of *Pichia pastoris* producing a fusion protein. Bioprocess Biosyst Eng 24:385–393.

Katakuga Y, Zhang W, Zhuang G, Omasa T, Kishimoto M, Goto Y, Suga KI. 1998. Effect of methanol concentration on the production of human beta-2-glycoprotein I domain V by a recombinant *Pichia pastoris:* A simple system for control of methanol concentration using a semiconductor gas sensor. J Ferment Bioeng 86:482–487.

Kobayashi K, Kuwae S, Ohya T, Ohda T, Ohyama M, Tomomitsu K. 2000. High level secretion of recombinant human serum albumin by fed-batch fermentation of the methylotrophic yeast, *Pichia pastoris*, based on optimal methanol feeding strategy. J Biosci Bioeng 90:280–288.

Kupcsulik B, Sevella B, Ballagi A, Kozma J. 2001. Evaluation of three methanol feed strategies for recombinant *Pichia pastoris* Muts fermentation. Acta Alimentaria 30:99–111.

Minning S, Serrano A, Ferrer P, Sola C, Schmid RD, Valero F. 2001. Optimization of high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris:* Enzyme production using plasmid pPICZ-alpha-A-ROL. J Biotechnol 86:59–70.

Murasugi A, Tohma-Aiba Y, Asami Y. 2000. Production of recombinant

human midkine in yeast, *Pichia pastoris:* Induction by methanol in high cell density fermentation. Biosci Bioeng 90:395–399.

O'Reilly MS, Boehm T, Shing Y, Fukai N, Vasios G, Lane WS, Flynn E, Birkhead JR, Olsen BR, Folkman J. 1997. Endostatin: An endogenous inhibitor of angiogenesis and tumor growth. Cell 88:277–285.

Tschopp JF, Brust PF, Cregg JM, Stillman CA, Gingeras TR. 1987. Expression of Lac Z gene from two methanol regulated promoters in *Pichia pastoris*. Nucleic Acid Res. 15:3859–3876.

Thommes J, Halfar M, Gieren H, Curvers S, Takors R, Brunschier R, Kula MR. 2001. Human chymotrypsinogen B production from *Pichia pastoris* by integrated development of fermentation and downstream processing. Part 2. Protein recovery. Biotechnol Prog 17:503–512.

Trinh LB, Noronha SB, Fannon M, Shiloach J. 2000. Recovery of mouse endostatin produced by *Pichia pastoris* using expanded bed adsorption. Bioseparation 9:223–230.

Wagner LW, Matheson NH, Heisey RF, Schineder K. 1997. Use of a silicone tubing sensor to control methanol concentration during fed batch fermentation of *P.pastoris*. Biotechnol Tech 11:791–795.

Zhang W, Bevins MA, Plantz BA, Smith LA, Meagher MM. 2000a. Modeling *Pichia pastoris* growth on methanol and optimizing the production of a recombinant protein, the heavy-chain fragment C of botulinum neurotoxin, serotype A. Biotechnol Bioeng 70:1–8.

Zhang W, Inan M, Meagher MM. 2000b. Fermentation strategies for recombinant protein expression in the methylotrophic yeast *Pichia pastoris*. Biotechnol Bioproc Eng 5:275–287

Ex. 1036 Bawa 2014

Bawa et al. Microbial Cell Factories 2014, 13:127
http://www.microbialcellfactories.com/content/13/1/127

MICROBIAL CELL FACTORIES

RESEARCH Open Access

Functional recombinant protein is present in the pre-induction phases of *Pichia pastoris* cultures when grown in bioreactors, but not shake-flasks

Zharain Bawa[1,2], Sarah J Routledge[1], Mohammed Jamshad[3], Michelle Clare[1], Debasmita Sarkar[1], Ian Dickerson[4], Markus Ganzlin[2,5], David R Poyner[1] and Roslyn M Bill[1*]

Abstract

Background: *Pichia pastoris* is a widely-used host for recombinant protein production; expression is typically driven by methanol-inducible alcohol oxidase (*AOX*) promoters. Recently this system has become an important source of recombinant G protein-coupled receptors (GPCRs) for structural biology and drug discovery. The influence of diverse culture parameters (such as pH, dissolved oxygen concentration, medium composition, antifoam concentration and culture temperature) on productivity has been investigated for a wide range of recombinant proteins in *P. pastoris*. In contrast, the impact of the pre-induction phases on yield has not been as closely studied. In this study, we examined the pre-induction phases of *P. pastoris* bioreactor cultivations producing three different recombinant proteins: the GPCR, human A_{2a} adenosine receptor ($hA_{2a}R$), green fluorescent protein (GFP) and human calcitonin gene-related peptide receptor component protein (as a GFP fusion protein; hCGRP-RCP-GFP).

Results: Functional $hA_{2a}R$ was detected in the pre-induction phases of a 1 L bioreactor cultivation of glycerol-grown *P. pastoris*. In a separate experiment, a glycerol-grown *P. pastoris* strain secreted soluble GFP prior to methanol addition. When glucose, which has been shown to repress *AOX* expression, was the pre-induction carbon source, $hA_{2a}R$ and GFP were still produced in the pre-induction phases. Both $hA_{2a}R$ and GFP were also produced in methanol-free cultivations; functional protein yields were maintained or increased after depletion of the carbon source. Analysis of the pre-induction phases of 10 L pilot scale cultivations also demonstrated that pre-induction yields were at least maintained after methanol induction, even in the presence of cytotoxic concentrations of methanol. Additional bioreactor data for hCGRP-RCP-GFP and shake-flask data for GFP, horseradish peroxidase (HRP), the human tetraspanins hCD81 and CD82, and the tight-junction protein human claudin-1, demonstrated that bioreactor but not shake-flask cultivations exhibit recombinant protein production in the pre-induction phases of *P. pastoris* cultures.

Conclusions: The production of recombinant $hA_{2a}R$, GFP and hCGRP-RCP-GFP can be detected in bioreactor cultivations prior to methanol induction, while this is not the case for shake-flask cultivations of GFP, HRP, hCD81, hCD82 and human claudin-1. This confirms earlier suggestions of leaky expression from *AOX* promoters, which we report here for both glycerol- and glucose-grown cells in bioreactor cultivations. These findings suggest that the productivity of *AOX*-dependent bioprocesses is not solely dependent on induction by methanol. We conclude that in order to maximize total yields, pre-induction phase cultivation conditions should be optimized, and that increased specific productivity may result in decreased biomass yields.

Keywords: *P. pastoris*, Methanol induction, Pre-induction expression, Bioprocess, Adenosine A_{2a} receptor, GPCR, GFP, CD81, CD82, Claudin-1, HRP, CGRP-RCP

* Correspondence: r.m.bill@aston.ac.uk
[1]School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, United Kingdom
Full list of author information is available at the end of the article

BioMed Central

© 2014 Bawa et al.; licensee BioMed Central. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Background

P. pastoris is a simple eukaryote with a fully sequenced genome [1,2], which can be cultured to high biomass yields without incurring high costs [3]. In its most widely used format, protein production is driven by a very strong, methanol-inducible alcohol oxidase (*AOX*) promoter, which is generally accepted to be tightly regulated [4]. To take full advantage of its attributes, *P. pastoris* must be cultured in controlled bioreactors, rather than shake-flasks, in cultivations that typically comprise several distinct phases. During the first or batch phase, cells grow at their maximum growth rate (μ_{max}) until the initial carbon source, typically glycerol, has been depleted. In a subsequent fed-batch phase, the same carbon source is fed continuously with the objective of yielding high pre-induction biomass; during this phase, growth is nutrient limited and a constant specific growth rate, lower than μ_{max}, is achieved. A transition phase, when the glycerol feed is stopped and the cells are monitored for glycerol depletion, allows the cells to adapt to low concentrations of inducer (typically methanol); in some cases temperature changes are also applied to facilitate induction at a temperature optimized for a given target protein. Finally, in the induction phase, methanol is added in a controlled manner [5,6] to induce *AOX*-driven recombinant protein production.

Recently *P. pastoris* has become an important source of recombinant G protein-coupled receptors (GPCRs) for structural biology and drug discovery. Drugs targeting GPCRs, which transmit a wide array of signals in different cell types, underpin modern medicine [7]; they account for about 40% of all prescription pharmaceuticals on the market. Relevant structural results are exemplified by the X-ray crystal structure of recombinant human histamine H1 receptor bound to doxepin at 3.1 Å resolution [8] and the 2.7 Å resolution structure of recombinant A_{2a} adenosine receptor ($hA_{2a}R$) in complex with an antibody Fab fragment that stabilizes an inactive form of the receptor [9].

Numerous studies have examined the influence of parameters such as the temperature and pH of the culture, the amount of dissolved oxygen (DO) in the culture medium and, specifically, the addition of chemical additives and ligands for optimal GPCR production in the pre-induction and induction phases of *P. pastoris* cultivations [4,10–13]. Notably, Singh and colleagues observed active $hA_{2a}R$ in bioreactor cultures prior to the methanol feed [14], while this was not apparent in shake flasks [15]. These findings are especially noteworthy because glycerol, glucose, ethanol and acetate have all been shown to support growth of *P. pastoris* cells without inducing the *AOX* promoter [16]. For example, Hellwig and colleagues [17] demonstrated glycerol in the culture medium inhibited production of a recombinant single-chain antibody in mixed feed bioreactor cultures, while Inan and Meagher reported that ethanol and acetate repress *AOX*-driven expression [16].

The first aim of our study was to specifically investigate the pre-induction phases of *P. pastoris* bioreactor cultivations producing recombinant $hA_{2a}R$. We examined a glycosylation-deficient $hA_{2a}R$ mutant designed by Fraser [18], who had previously achieved a functional yield of 4 pmol mg^{-1} in shake flask cultures with no pre-induction activity. Prior to methanol induction, we detected baseline production of $hA_{2a}R$ confirming that expression driven by the *AOX* promoter is leaky in a glycerol-containing medium. In these cultures the baseline yield of $hA_{2a}R$ was not substantially improved in the methanol induction phase and pre-induction yields of $hA_{2a}R$ were maintained even at cytotoxic levels of methanol. In order to assess whether these observations were specific to $hA_{2a}R$ as the recombinant target protein, two additional strains were investigated; one that secretes the model protein, soluble green fluorescent protein (GFP) and one that secretes the peripheral membrane protein, human calcitonin gene-related peptide receptor component protein (as a fusion protein with GFP; hCGRP-RCP-GFP). When grown on glycerol or glucose as the pre-induction carbon source, GFP (as for $hA_{2a}R$) was produced in the pre-induction phases. Following methanol induction, yields were not substantially improved for glycerol-grown cells, but were increased by a factor of 4 without a corresponding increase in biomass for glucose-grown cells. Pre-induction expression of hCGRP-RCP-GFP was detected in glycerol-grown cells. In contrast, there was no pre-induction expression in shake-flask cultures of strains producing GFP, horseradish peroxidase (HRP), the human tetraspanins hCD81 and CD82 or the tight-junction protein human claudin-1. These data demonstrate that bioreactor but not shake-flask cultivations exhibit recombinant protein production in both glycerol- and glucose-grown pre-induction phases. This highlights the importance of understanding the influence of pre-induction phase cultivation conditions on both recombinant protein yields and *P. pastoris* biomass yields.

Results

$hA_{2a}R$ is produced in the pre-induction phases of glycerol-grown *P. pastoris* bioreactor cultivations

Duplicate 1 L bioreactor cultivations of *P. pastoris* producing a glycosylation-deficient $hA_{2a}R$ mutant [18] were examined over a 91 h cultivation period. Since it is established that lowering the temperature to 22°C during induction increases the functional yield of $hA_{2a}R$ [13,14,18], the cultivation temperature was lowered during phase III. Figure 1 shows the glycerol batch (I), glycerol fed-batch (II), transition (III) and methanol induction (IV) phases analysed in this study together with the residual glycerol measurements, dry cell weight (DCW) and B_{max} estimates for recombinant $hA_{2a}R$ binding activity, as measured

Motif Exhibit 1059, 417 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

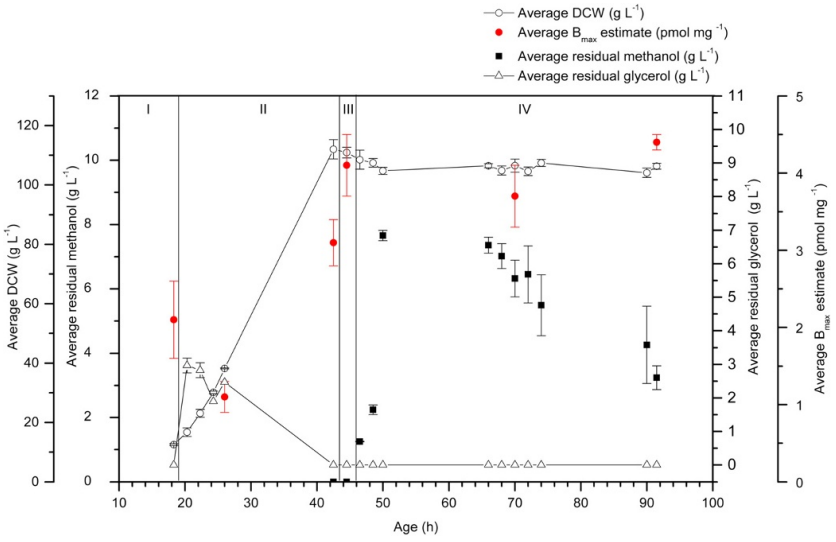


Figure 1 $hA_{2a}R$ binding activity is present in all phases of 1 L glycerol-grown *P. pastoris* bioreactor cultivations. Duplicate 1 L bioprocesses were analysed for $hA_{2a}R$ binding activity (pmol mg^{-1}; red circles), DCW (g L^{-1}; white circles), residual glycerol concentration (g L^{-1}; white triangles) and residual methanol concentration (g L^{-1}; black squares). $hA_{2a}R$ binding activity was measured in all the phases, including pre-induction phases I and II. The residual glycerol concentration was at its highest during phase II, dropping to zero in phase II and for the duration of the cultivation. DCW increased during phases I and II and plateaued during phases III and IV. Residual methanol ranged from 1.25 – 7.66 g L^{-1} during the induction phase. Measurements were made in triplicate for each culture. The error bars represent standard error of the mean (SEM).

by radio-ligand binding with the tritiated antagonist, ZM241385. Residual glycerol peaked during phase II at ~ 2 g L^{-1} and was not present during the transition and induction phases (III and IV). Binding activity between 1.1 and 3.1 pmol mg^{-1} was measured during the batch and fed-batch phases (I and II), indicating leaky expression. During the transition phase (III), binding activity increased to 4.1 pmol mg^{-1} and during the induction phase (IV) it was 3.7 – 4.4 pmol mg^{-1}. The specific yield from the 1 L cultivation was 122.2 pmol g^{-1} and the total yield was 12,986.3 pmol. Notably, the impact of the methanol feed was minimal since the pre-induction specific yield was not substantially increased (Figure 1).

$hA_{2a}R$ is produced in glycerol-grown *P. pastoris* cultivations in the absence of a methanol induction step, while pre-induction binding activity is maintained at pilot scale, even at cytotoxic methanol concentrations

$hA_{2a}R$ was also produced in all phases of a completely methanol-free cultivation of glycerol-grown *P. pastoris* (Figure 2). During the fed-batch phase (II), when glycerol was present at 1.9 g L^{-1}, the recombinant $hA_{2a}R$ yield was 1.1 pmol mg^{-1}, reaching a final yield of 1.6 pmol mg^{-1} at the end of the cultivation. The specific yield from the 1 L cultivation was 90.3 pmol g^{-1} and the total yield was 5,598.3 pmol. This was lower than the yield achieved in the corresponding induced culture (Figure 1; 122.2 pmol g^{-1} and 12,986.3 pmol at the end of phase IV) indicating the positive impact of the methanol feed on total yield. The DCW reached a maximum of 62.0 g L^{-1} in contrast to that of 106.3 g L^{-1} for the corresponding induced culture (Figure 1).

To examine whether these findings transferred to larger vessels, two simultaneous pilot scale bioreactor cultivations (10 L starting volume) of *P. pastoris* expressing $hA_{2a}R$ were compared following the same batch, fed-batch and transition feeding regimes (Table 1). During the induction phase, different methanol feed profiles were applied: a low methanol feed profile with methanol concentrations maintained below 5.0 g L^{-1} and a higher methanol feed profile with methanol concentrations allowed to exceed 5.0 g L^{-1}. A total cultivation time of 91.2 h included the following phases: batch (phase I), fed-batch (phase II), transition (phases IIIA and IIIB) and induction (phase IV and V). An identically-conducted batch phase (I) on 10 g L^{-1} glycerol lasted for 19.1 h for both bioprocesses. At this point, the characteristic DO spike indicated full consumption of the glycerol carbon source and the stirrer speed decreased as a result of the DO control. In the subsequent 20 h fed-batch phase (II), the same exponential, growth rate-limiting glycerol feed profile of 4 g L^{-1} h^{-1} was applied to both bioprocesses, and increased exponentially at a rate of 0.15 h^{-1} for 10 h and then 0.03 h^{-1} until the end of this phase. The transition phase had two different sections (IIIA and IIIB): in phase IIIA, no feed was applied to the cultivations for 1 h; in phase IIIB a constant methanol feed

Motif Exhibit 1059, 418 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

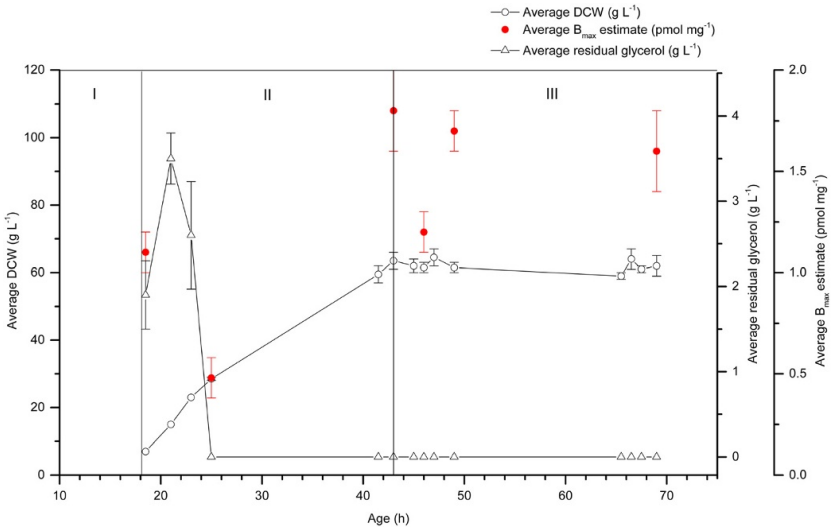


Figure 2 $hA_{2a}R$ binding activity is present in all phases of 1 L methanol-free cultivations. Duplicate 1 L bioprocesses for the same strain cultured previously (Figure 1) were analysed for $hA_{2a}R$ binding activity (pmol mg^{-1}; red circles), DCW (g; white circles) and residual glycerol concentration (g L^{-1}; white triangles) in the absence of methanol induction. $hA_{2a}R$ binding activity was measured in all the phases: glycerol batch phase (I); glycerol fed-batch phase (II) and carbon source starvation phase (III). The residual glycerol concentration was at its highest during phase I, dropping to zero in phase II and for the duration of the cultivation. DCW increased during phases I and II and plateaued during phases III. Measurements were made in triplicate. The error bars represent the SEM.

of 4 g L^{-1} h^{-1} was applied. At the end of phase IIIB (47 h), the cultivation temperature was lowered to 22°C and then the induction phase commenced differently for the two cultivations. For both cultivations, methanol was added in an exponential manner. For the low methanol fed cultivation, a total of 2.2 kg of methanol was added and for the high methanol fed cultivation, a total of 3.8 kg methanol was added by the end of the cultivation.

Table 1 shows B_{max} estimates, residual glycerol and methanol concentrations and the amount of biomass generated (DCW and total membrane protein) for both cultivations during the transition and induction phases. The low methanol cultivation produced the highest total and specific yields of $hA_{2a}R$. Even in the absence of methanol, $hA_{2a}R$ binding activity was 4.4 pmol mg^{-1} for both the low (SEM = ± 0.1 pmol mg^{-1}) and high methanol

Table 1 Pre-induction yields of $hA_{2a}R$ are maintained at pilot scale even in the presence of cytotoxic methanol concentrations

	Phase	Age (h)	Residual glycerol (g L^{-1})	Residual methanol (g L^{-1})	DCW (g L^{-1})	Total membrane protein yield in 1 L culture (mg)	B_{max} estimate (pmol mg^{-1})	Total $hA_{2a}R$ yield in 1 L culture (pmol)	Specific yield (pmol g^{-1})
Low methanol cultivation	IIIA	42.2	0 (0)	0 (0)	41.91 (0.6)	764.4	4.4 (0.1)	3363.6	80.3
	IIIB	44.4	1.1 (0.2)	0.5 (0.1)	42.10 (0.4)	2791.1	7.1 (0.6)	19816.9	470.7
	IV	66.4	0 (0)	1.0 (0.1)	50.53 (0.5)	4084.4	4.5 (0.2)	18380.0	363.7
	IV	89.0	0 (0)	0 (0)	71.82 (0.2)	7704.4	5.0 (0.2)	38522.2	536.4
High methanol cultivation	IIIA	42.2	2.9 (0.1)	0 (0)	46.71 (0.5)	813.3	4.4 (0.2)	3578.7	76.6
	IIIB	44.4	4.2 (0.2)	2.1 (0.2)	46.66 (0.7)	3060.0	4.0 (0.0)	12240.0	262.3
	IV	66.4	1.1 (0.2)	39.1 (9.0)	45.76 (0.5)	3277.8	3.8 (0.1)	12455.5	272.2
	V	88.9	0 (0)	84.7 (0.5)	37.64 (0.8)	1797.8	3.1 (0.1)	5573.1	148.1

The total yield of $hA_{2a}R$ (pmol) and the specific yield of $hA_{2a}R$ per gram of DCW (pmol g^{-1}) were derived from the B_{max} and total membrane protein measurements, as described in the Methods section. Values are calculated per L of culture for the transition and induction phases of the low and high methanol cultivations and are the mean of triplicate determinations, with the standard error of the mean in parentheses. Triplicate measurements were from three different membrane preparations and standard error of the mean is shown in parentheses.

Motif Exhibit 1059, 419 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

(SEM = ± 0.2 pmol mg^{-1}) cultivations (Table 1). For the low methanol cultivation, in phase IIIB, where a constant methanol feed was initiated and glycerol was still present at 1.1 g L^{-1}, functional receptor was detected. In phases IV and V of the low methanol cultivation, no glycerol was present and binding activity was present. For the high methanol cultivation during the transition and induction phases (IIIB and IV), glycerol was present in the culture medium and binding activities were 4.0 and 3.8 pmol mg^{-1}, respectively. In phase V, there was no residual glycerol in the culture medium and binding activity was 3.1 pmol mg^{-1}.

GFP is produced in the pre-induction phases of glycerol-grown *P. pastoris* bioreactor cultivations

In order to assess whether pre-induction expression was specific to $hA_{2a}R$ as the recombinant target protein, we examined the production of recombinant soluble GFP. Figure 3 shows the data for duplicate 1 L bioreactor cultivations with glycerol as the pre-induction carbon source. The residual glycerol concentration was at its highest during phase I, eventually dropping to zero in phase II: during the batch phase (I), glycerol was present at 3.0 g L^{-1} and the recombinant GFP yield was 2.6 mg L^{-1} prior to methanol addition. The GFP yield increased from 2.7 to 3.5 mg L^{-1} in the fed-batch phase (II; the glycerol was consumed 3 h into this fed-batch phase), reaching a final yield of 4.7 mg L^{-1} in the induction phase (IV). Residual methanol ranged from 3.01 to 8.52 g L^{-1} and DCW reached a maximum of 175 g L^{-1}.

$hA_{2a}R$ and GFP are produced in the pre-induction phases of glucose-grown *P. pastoris* bioreactor cultivations

While glycerol de-represses *AOX* expression, glucose has been shown to repress *AOX* expression [19] even in the presence of methanol [20]. Duplicate cultivations were therefore analysed in which glucose replaced glycerol as the pre-induction carbon source in the recombinant production of both $hA_{2a}R$ (Figure 4) and GFP (Figure 5). For the production of $hA_{2a}R$ (Figure 4), the residual glucose concentration ranged from 0.01 to 0.04 mM; glucose was not fully consumed by the cells during the cultivation. However, when the residual glucose was very low during the fed-batch phase (phase II; 0.012 mM), the B_{max} was 1.2 pmol mg^{-1} which was similar to B_{max} values during methanol induction (phase IV; 1.17 to 1.46 pmol mg^{-1}). The residual glucose also remained above zero during the transition phase (III, no glucose added). The $hA_{2a}R$ yield decreased between 26 and 46 hours (phase II and III) and then increased to 1.17 – 1.46 pmol mg^{-1} when methanol was added for the induction phase (IV). Methanol levels reached cytotoxic levels rapidly (99.76 g L^{-1}) by the end of the cultivation which enabled us to confirm our earlier findings at pilot scale for glycerol-grown cultures (Table 1): we observed that for the glucose-grown cultures in Figure 4, pre-induction yields could be at least maintained after methanol induction, even in the presence of

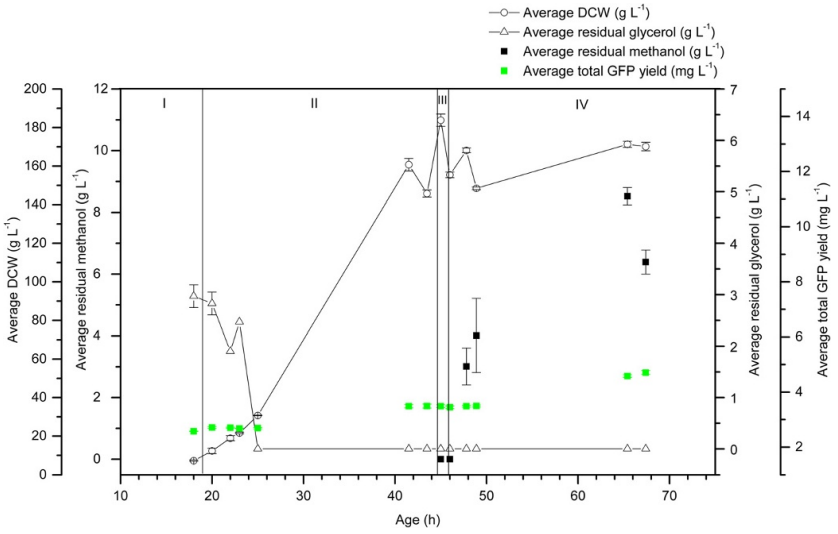


Figure 3 Recombinant GFP is produced in the pre-induction phases of 1 L glycerol-grown *P. pastoris* bioreactor cultivations. GFP yield (mg L^{-1}; green squares) was measured in the culture supernatant in all phases of duplicate 1 L glycerol-grown *P. pastoris* cultures. The residual glycerol concentration (g L^{-1}; white triangles) was at its highest during phase I, eventually dropping to zero in phase II. The DCW (white circles) for this cultivation reached a maximum of 175 g L^{-1}. Residual methanol (g L^{-1}; black squares) ranged from 3.01 to 8.52 g L^{-1}. All measurements were made in triplicate and error bars represent the SEM.

Motif Exhibit 1059, 420 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

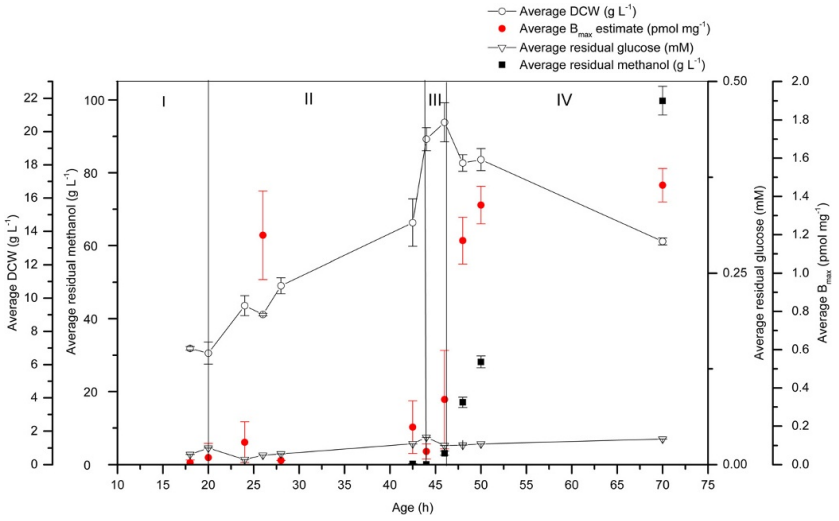


Figure 4 Recombinant $hA_{2a}R$ is produced in the pre-induction phases of 1 L glucose-grown *P. pastoris* bioreactor cultivations. $hA_{2a}R$ binding activity (pmol mg^{-1}; red circles) was measured in the culture supernatant in all the phases of a 1L glucose-grown *P. pastoris* cultivation. The residual glucose concentration (mM; white inverted triangles) ranged from 0.01-0.04 mM during the cultivation was not consumed completely by the cells. The residual methanol (g L^{-1}; black squares) ranged from 0.20 to 99.76 g L^{-1}. The DCW (g L^{-1}; white circles) for this cultivation reached a maximum of 21 g L^{-1}. All measurements were made in triplicate and error bars represent the SEM.

cytotoxic concentrations of methanol. The DCW reached a maximum of 21 g L^{-1} by the end of the cultivation. These results differed from the glycerol grown cells (Figure 1): the DCW weight produced by the glucose-grown cultivation was only 21 g L^{-1} when compared to greater than 100 g L^{-1} for the glycerol-grown cultivation. The $hA_{2a}R$ levels also differed: better yields were produced from the glycerol-grown cultivation (a maximum of 4.4 pmol mg^{-1} compared to a maximum of 1.46 pmol mg^{-1} at cytotoxic methanol levels, respectively).

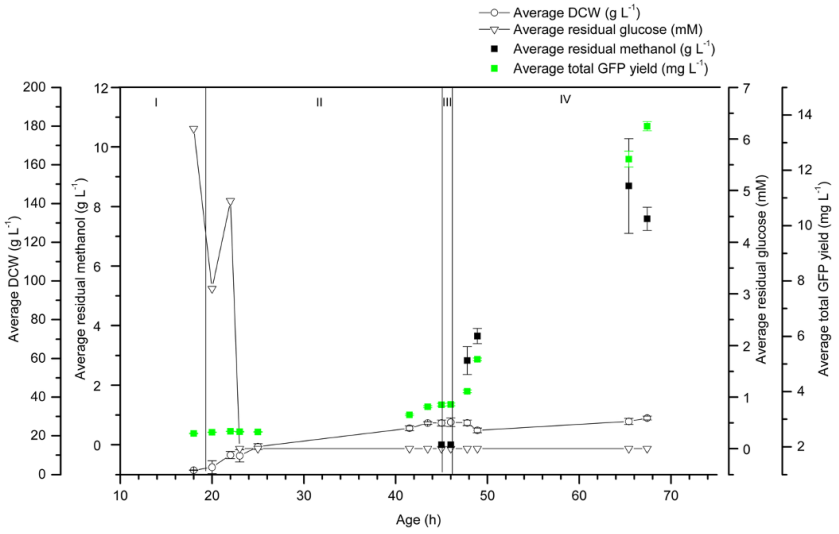


Figure 5 Recombinant GFP is produced in the pre-induction phases of 1 L glucose-grown *P. pastoris* bioreactor cultivations. GFP yield (mg L^{-1}; green squares) was measured in the culture supernatant in all phases of duplicate 1 L glucose-grown *P. pastoris* cultivations. The residual glucose concentration (mM; white inverted triangles) was at its highest during phase I, eventually dropping to zero in phase II. The residual methanol (g L^{-1}; black squares) ranged from 2.83 to 8.69 g L^{-1}. The DCW (g L^{-1}; triangles) for this cultivation reached a maximum of 30 g L^{-1}. All measurements were made in triplicate. The error bars represent the SEM.

Motif Exhibit 1059, 421 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

For the production of GFP (Figure 5), the residual glucose concentration was at its highest during phase I, eventually dropping to zero in phase II: during the batch phase (I), when the glucose concentration was 6 mM, the yield of recombinant GFP was 2.5 mg L^{-1} prior to methanol addition. This increased from 2.5 to 3.5 mg L^{-1} in the fed-batch phase (II; the glucose was consumed 3 h into this fed-batch phase), reached a plateau in the transition phase (III) and then increased during the induction phase (IV) to a maximum value of 13.6 mg L^{-1}. The residual methanol concentration ranged from 2.83 to 8.69 g L^{-1} and the DCW reached a maximum of 30 g L^{-1}. Once the pre-induction carbon source was depleted, DCW values for glycerol-grown cells (Figure 3) were approximately 5 times higher than for glucose-grown cells in the equivalent phase. However, glucose-grown cells exhibited a higher specific productivity than glycerol-grown cells, producing 3 times the yield of GFP at the expense of their biomass yields.

GFP is produced in both glycerol- and glucose- grown *P. pastoris* cultivations in the absence of a methanol induction step

Since GFP was present in the pre-induction phases of both glycerol- and glucose-grown *P. pastoris* bioreactor cultivations (Figures 3 and 5), its production was analyzed in the absence of a methanol induction step (Figure 6). GFP yield was measured in the culture supernatant in all phases of two 1 L *P. pastoris* cultivations, one grown on glycerol and one grown on glucose (Figure 6). In these cultivations, the transition phase (III) was extended from 2 h to 30 h. During the batch phase (I) of the glycerol-grown culture (Figure 6; upper graph), glycerol was present at 2.9 g L^{-1} and the recombinant GFP yield was 2.1 mg L^{-1} in line with the earlier glycerol-grown cultivations (2.6 mg L^{-1} prior to methanol addition, Figure 3). The total GFP yield remained stable in the fed-batch phase (II) and reached a final yield of 8.4 mg L^{-1} at the end of the cultivation. This was higher than the yield achieved in the corresponding induced cultures (4.7 mg L^{-1} at the end of phase IV; Figure 3) indicating the negative impact of the methanol feed on total yield. The DCW for the methanol free cultivation reached a maximum of 92.5 g L^{-1} (Figure 6; upper graph) in contrast to that of 175 g L^{-1} for the corresponding induced culture (Figure 3).

During the batch phase (I) of the glucose-grown culture (Figure 6; lower graph), glucose was present at 3.4 mM and the recombinant GFP yield was 1.8 mg L^{-1}, which was lower than that for the earlier glucose-grown cultivation (2.5 mg L^{-1}; Figure 3). The total GFP yield increased in the fed-batch phase (II), remained stable in the transition phase and reached a final yield of 5.6 mg L^{-1} at the end of the cultivation. This was lower than the yield achieved in the corresponding induced culture (13.6 mg L^{-1} at the end of phase IV; Figure 3) indicating the positive impact of the methanol feed on total yield. The DCW reached a maximum of 46.5 g L^{-1} in contrast to that of 30 g L^{-1} for the corresponding induced culture (Figure 5).

hCGRP-RCP-GFP is also produced in pre-induction phases of glycerol-grown *P. pastoris* bioreactor cultivations, while pre-induction yields of five proteins are not detected in shake flask cultivations

To examine whether these findings were applicable to other target proteins, we examined the pre-induction phases of glycerol-grown bioreactor cultivations producing hCGRP-RCP-GFP [21]: hCGRP-RCP-GFP was detected in the pre-induction phases (Table 2, Additional file 1). In contrast, shake-flask cultivations of GFP [22], HRP, hCD81 [23], hCD82 [23] and human claudin-1 [23] showed no recombinant protein production in the pre-induction phases (Table 2, Additional file 1).

Discussion

It is established that both the concentration and rate of addition of methanol to *P. pastoris* cultures can substantially affect recombinant protein yields [20,24–28]. In contrast, the impact of the pre-induction phases of *P. pastoris* cultures on yield has not been examined systematically in either bioreactor or shake-flask formats.

A significant finding emerging from this work is that an integral membrane protein ($hA_{2a}R$), a soluble protein (GFP) and a peripheral membrane protein (hCGRP-RCP-GFP) are produced prior to induction with methanol in bioreactor cultivations of *P. pastoris*. This is consistent with a previous report by Singh and colleagues who noted pre-induction $hA_{2a}R$ binding activity in glycerol-grown bioreactor cultivations of *P. pastoris*, an observation that was not further elaborated upon [14]. In a different bioreactor study by Çelik and colleagues, recombinant hEPO appeared to be produced prior to the onset of the methanol feed, but this was not discussed by the authors [24]. In contrast, pre-induction $hA_{2a}R$ binding activity was not apparent in shake-flasks using complex medium in studies by Singh and colleagues [15] or by Fraser [18]. In agreement with this, we show that GFP, HRP, hCD81, hCD82 and human claudin-1 are not detected in the pre-induction phases of shake-flask cultivations. The reason for this is currently unknown; we suspect it may be a function of the different oxygenation of cultures in the two formats and/or the different media used (complex in shake-flasks and minimal in bioreactors) [29].

The demonstration of recombinant protein synthesis prior to induction is especially noteworthy because glycerol, glucose, ethanol and acetate have all been shown to support growth of *P. pastoris* cells without inducing the *AOX* promoter [19]; our systematic study demonstrates that recombinant $hA_{2a}R$ and GFP can be detected

Motif Exhibit 1059, 422 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

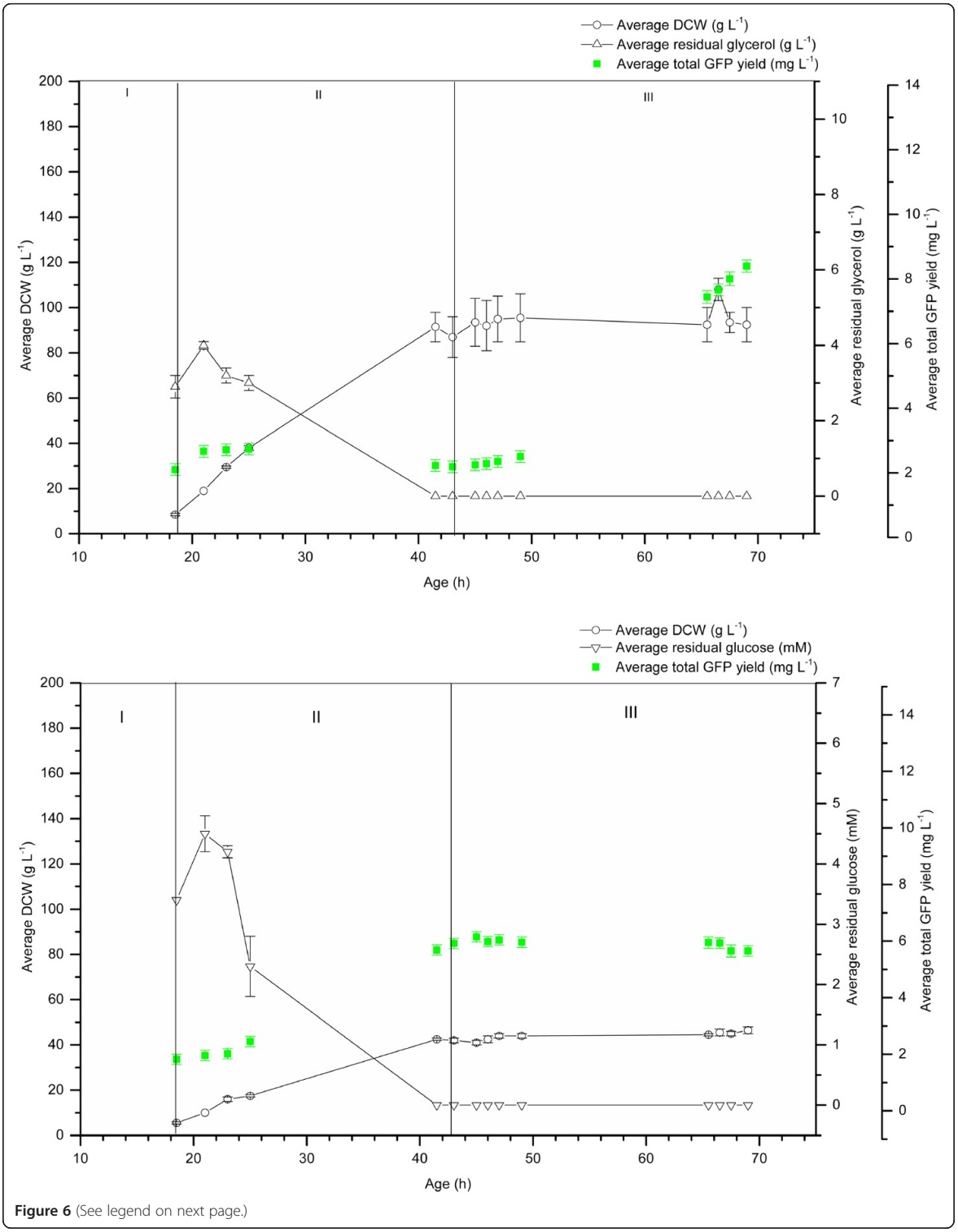


Figure 6 (See legend on next page.)

Motif Exhibit 1059, 423 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

(See figure on previous page.)
Figure 6 Recombinant GFP is produced in methanol-free cultivations of *P. pastoris* grown on either glycerol or glucose as the sole carbon source. GFP yield (mg L^{-1}; green squares) was measured in the culture supernatant in all phases of two 1 L *P. pastoris* cultivations, one grown on glycerol (upper graph) and one grown on glucose (lower graph). The residual glycerol concentration (upper graph; g L^{-1}; white triangles) was at its highest during phase I, eventually dropping to zero in phase II. The DCW (upper graph; g L^{-1}; white circles) for this cultivation reached a maximum of 92.5 g L^{-1}. The residual glucose concentration (lower graph; mM; white inverted triangles) eventually dropped to zero in phase II. The DCW (lower graph; g L^{-1}; white circles) for this cultivation reached a maximum of 46.5 g L^{-1}. All measurements were made in triplicate the error bars represent the SEM.

in both glycerol- and glucose-grown bioreactor cultivations prior to methanol induction, confirming earlier suggestions of leaky expression from *AOX* promoters. Our data therefore suggest that the productivity of *AOX*-dependent bioprocesses is not solely reliant on induction by methanol and that in order to maximize total yields, the impact on specific productivity of pre-induction phase cultivation should be optimized.

Our data demonstrate that methanol induction does not necessarily increase the pre-induction specific yield, measured in the transition phase, for cells initially cultured on glycerol. The total yield of $hA_{2a}R$ was at least maintained once the methanol feed commenced during the induction phase (Figure 1, Table 1), but the increase in total yield was not substantial, being only approximately double that of the pre-induction yield (Figure 1, Table 1). A subsequent glycerol-grown *P. pastoris* cultivation also produced recombinant GFP in the pre-induction phases, with similar results following induction with methanol (Figure 3). These increases in yield are similar to those seen by Çelik and colleagues in their hEPO cultivations [24].

Pre-induction $hA_{2a}R$ yields were maintained even under cytotoxic methanol concentrations (Table 1 and Figure 4). Cunha and colleagues [30] and Schenk and colleagues [31] demonstrated that residual methanol concentrations between 2 g L^{-1} and 3.5 g L^{-1} were optimal for soluble protein production, while methanol concentrations between 3.7 g L^{-1} and 20 g L^{-1} were cytotoxic and inhibited growth [30,31], consistent with our findings. We previously reported that residual methanol levels below 2.98 g L^{-1} gave optimal yields of recombinant GFP (3.74 g) when following a mixed 60% methanol and 40% sorbitol induction regime. In contrast, a 100% methanol induction regime, where the residual methanol reached a maximum level of 182.25 g L^{-1}, gave a poor recombinant GFP yield (0.09 g) [27].

Previously, we have demonstrated that high yields of recombinant GFP (3.74 g) but low biomass yields (OD_{595} = 64) were produced when *P. pastoris* was grown at a low induction phase growth rate (μ = 0.006 h^{-1}). In contrast, a higher induction phase growth rate (μ = 0.015 h^{-1}) resulted in a higher biomass yield (OD_{595} = 74), but a lower total GFP yield (0.98 g) [27]. Çalık and colleagues [32] applied three different growth rates in the methanol induction phase of *P. pastoris* cultures expressing soluble human growth hormone: 0.02, 0.03 and 0.04 h^{-1} using a mixed sorbitol and methanol feed. The highest biomass yield (48 g L^{-1}) was obtained at a growth rate of 0.04 h^{-1}, but the highest yield of recombinant human growth hormone (270 mg L^{-1}) was produced at 0.03 h^{-1}. The authors suggested that this might be due to the lower growth rate producing lower biomass and therefore fewer extracellular

Table 2 Functional recombinant protein is present in the pre-induction phases of *P. pastoris* cultures when cultivated in bioreactors, but not in shake-flasks

Protein	Bioreactor		Shake-flask	
	Pre-induction expression in glycerol-grown cells?	*Pre-induction expression in glucose-grown cells?*	*Pre-induction expression in glycerol-grown cells?*	*Pre-induction expression in glucose-grown cells?*
Integral membrane proteins				
$hA_{2a}R$	Y (Figures 1 and 2)	Y (Figure 4)	N [18]	–
CD81	–	–	N (Additional file 1: Figure S1)	–
CD82	–	–	N (Additional file 1: Figure S1)	N (Additional file 1: Figure S1)
Claudin-1	–	–	N (Additional file 1: Figure S1)	N (Additional file 1: Figure S1)
Peripheral membrane protein				
hCGRP-RCP-GFP	Y (Table S1)	–	–	–
Soluble proteins				
GFP	Y (Figures 3 and 6A)	Y (Figures 5 and 6B)	N (Additional file 1: Table S2)	N (Additional file 1: Table S2)
HRP	–	–	N (Additional file 1: Table S3)	–

The presence (Y) or absence (N) of recombinant protein was determined in bioreactor and shake-flask cultivations for glycerol- and glucose-grown cells as indicated; the evidence supporting these designations is given in parentheses and refers to figures within the text or in supplementary files; "–" indicates that the experiment was not done.

Motif Exhibit 1059, 424 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

proteases being present in the cultivation. Similar trends were also observed for GFP, hEPO, human serum albumin, heavy-chain fragment C of *Botulinum* neurotoxin serotype A and avidin [24,26,27,33].

Glucose has previously been shown to support growth of *P. pastoris*, but at a lower specific growth rate and yielding lower amounts of biomass than that achieved by cells grown on glycerol [34]. Following induction with methanol, DCW values for glycerol-grown cells producing GFP were approximately 5 times higher than for glucose-grown cells in the equivalent pre-induction phase (Figures 3 and 5), although the glucose-grown cells produced 3 times the GFP yield of the glycerol-grown cells (Figures 3 and 5). Notably, the methanol feed had a positive impact on yield for glucose-grown cells, but a negative impact on yield for glycerol-grown cells (Figures 3, 5 and 6). For $hA_{2a}R$ produced in glucose-grown cultures (Figure 4), methanol also had a positive impact on yield. However, the yields compared to the glycerol-grown cells were lower, which was in contrast to GFP yields produced in glucose-grown cells.

At pilot scale, the methanol feed influenced the biomass yield and particularly increased the total $hA_{2a}R$ yield in the low methanol cultivation compared to the high methanol cultivation (Table 1). Notably, a similar increase in total membrane protein yield has previously been reported to increase total volumetric yields of GPCRs in a respiratory yeast strain of *Saccharomyces cerevisiae* [35]. As a consequence of increased total membrane protein yield, the yield per unit volume of the low methanol cultivation was approximately 40 times higher than shake flask cultivations of the same $hA_{2a}R$ construct [18]; total membrane protein yields from shake flasks are 200 mg [18] compared with almost 8,000 mg in bioreactors (Table 1). This yield improvement compares very favourably with that achieved by Singh and colleagues using an optimized $hA_{2a}R$ construct under broadly similar bioprocess conditions to our study [15]: on transferring from shake flasks to bioreactors a yield improvement of approximately 25 times was achieved on account of both an increase in specific productivity (25 pmol mg^{-1} in a shake flask; 100 pmol mg^{-1} in a bioreactor) and biomass yield (OD_{600} = 13 in a shake flask; OD_{600} = 80 in a bioreactor). Overall, our data suggest that specific productivity may not necessarily be solely dependent on methanol induction and can be increased at the expense of improvements in biomass yields in the methanol induction phase.

Conclusions

Production of $hA_{2a}R$, GFP and hCGRP-RCP-GFP was detected in bioreactors prior to methanol induction in glycerol- and glucose-grown cells; in contrast, GFP, HRP, hCD81, hCD82 and human claudin-1 were not detected in the pre-induction phases of shake-flask cultivations. In bioreactors, the transition phase yield was not necessarily increased following methanol induction and was maintained even at cytotoxic methanol concentrations. Using glucose as a pre-induction carbon source, specific productivity could be increased at the expense of improvements in biomass yield, with methanol having a positive impact on these cells. These data provide a platform to further optimize the production of recombinant proteins, and especially GPCRs [7] as part of modern drug discovery programmes.

Methods

Yeast strains and culturing conditions

P. pastoris strains X33 $hA_{2a}R$ [18], X33 GFP [22,29], SMD1163 hCGRP-RCP-GFP [21], CBS 7435 Mut^s HRP, X33 hCD81 [23], X33 hCD82 [23] and X33 human claudin-1 [23] were used in all experimental procedures using previously-described culture conditions. To generate a seed culture for the bioreactor studies or to establish a shake-flask culture, a single, freshly-streaked colony was used to inoculate 5 mL YPD medium (2% peptone, 1% yeast extract, 2% glucose), and this was incubated at 30°C in a shaking incubator for 2 days. The 5 mL YPD pre-culture was then used to inoculate 50 mL YPD in a 250 mL baffled shake-flask, and this was incubated at 30°C in a shaking incubator overnight to an OD_{600} of 15. For bioreactor cultivations, these cells were then used to inoculate 200 mL FM22 minimal medium, which was composed of (g L^{-1}) KH_2PO_4, 42.9; $(NH_4)_2SO_4$, 5; $CaSO_4{\cdot}2H_2O$, 1.0; K_2SO_4, 14.3; $MgSO_4{\cdot}7H_2O$, 11.7; and glycerol, 10 with *Pichia* trace minerals 4 (PTM_4) salt solution composed of (g L^{-1}) $CuSO_4{\cdot}5H_2O$, 2.0; NaI,0.08; $MnSO_4{\cdot}H_2O$, 3.0; $Na_2MoO_4{\cdot}2H_2O$, 0.2; H_3BO_3, 0.02; $CaSO_4{\cdot}2H_2O$, 0.5; $CoCl_2$, 0.5; $ZnCl_2$, 7; $FeSO_4{\cdot}7H_2O$, 22; biotin, 0.2; 1 ml of concentrated H_2SO_4 . These bioreactor seeder flasks were incubated at 30°C and 220 rpm overnight to an OD_{600} of 30.

1 L bioreactor cultures

For the glycerol cultivations, 1 L FM22 medium supplemented with 0.5 mL J673A (Struktol), 4 mL PTM_4 salts, 0.8 mL 10 g L^{-1} biotin and 10 g 100% glycerol was transferred to a 2 L (total volume) jacketed glass bioreactor (Applikon Biotechnology Ltd). Culture temperature was maintained at 30°C and pH at 5. The maximum agitation rate was 1,250 rpm. The air flow rate was 2 L min^{-1} and the DO set point was 30%. The end of the glycerol batch phase was indicated by a spike in DO to 100%. A glycerol (50% v/v aqueous glycerol) fed-batch phase was maintained for 40 h at a flow rate of 14 mL min^{-1}. For the methanol induced cultivations, the transition phase was maintained for 2 h and the temperature was lowered to 22°C for cultivations producing $hA_{2a}R$ but remained at 30°C for the GFP cultivations. The induction phase

Motif Exhibit 1059, 425 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

comprised feeding 50% (v/v) aqueous methanol at 2.04 mL min^{-1} for 40 h. For the glucose cultivations, 1 L FM22 medium supplemented with 0.5 mL J673A (Struktol), 4 mL PTM_4 salts, 0.8 mL 10 g L^{-1} biotin and 10% glucose was transferred to a 2 L (total volume) jacketed glass bioreactor (Applikon Biotechnology Ltd). Culture temperature was maintained at 30°C and pH at 5. The maximum agitation rate was 1,250 rpm. The end of the glucose batch phase was indicated by a spike in DO to 100%. A glucose (10%) fed-batch phase was maintained for 40 h at a flow rate of 14 mL min^{-1}. For the methanol-free cultivations, the conditions remained the same as in the methanol induced cultivations for the batch and fed-batch phases. At the transition phase, the glycerol or glucose feed was stopped, the temperatures changed (22°C for the $hA_{2a}R$ cultivation and 30°C for the GFP cultivation) and the cultivations were allowed to run for a further 26 h without any methanol feed or any other carbon source. Bioreactor control was via an Applikon ADI1010 control unit.

10 L bioreactor cultures

10 L FM22 medium supplemented with 1 mL P2000 (Sigma-Aldrich), 40 mL PTM_4 salts and 8 mL 10 g L^{-1} biotin were transferred into each of two 30 L (total volume) jacketed steel bioreactors (Biostat C, Sartorius Ltd). Culture temperature was maintained at 30°C and pH at 5. The maximum agitation rate was 1,500 rpm. For both cultivations, the end of the glycerol batch phase (phase I) was indicated by a spike in DO to 100%. A glycerol fed-batch phase (phase II) was then maintained for 40 h by employing an exponential feed rate of 50% aqueous glycerol (v/v) at 4 g L^{-1} h^{-1} increasing exponentially first at a rate of 0.15 h^{-1} for 10 h and subsequently at a rate of 0.03 h^{-1}. The first section of the transition phase (phase IIIA; starvation), during which no further carbon source was fed into the bioreactor, was maintained for 1 h. The second section of the transition phase (phase IIIB) comprised a further 1 h, where a constant methanol feedstock (50% (v/v) aqueous methanol) was applied at 8 g L^{-1} h^{-1}. The culture temperature was lowered to 22°C and an exponential methanol (50% (v/v) aqueous methanol) feed profile was applied by exponentially increasing the feed rate at 0.01 h^{-1} (low methanol cultivation) or 0.03 h^{-1} (high methanol cultivation). The induction phase continued for 40 h.

Sampling

Samples (1 mL) were taken for optical density measurements at 600 nm (OD_{600}). For dry cell weight measurements, 1.5 mL culture was sampled in triplicate, placed in pre-weighed tubes and centrifuged at 5,000 × g for 5 min. The supernatant was removed and stored at −20°C for residual glycerol and/or residual methanol analysis and the tubes were placed in a 100°C oven with the lids open and dried overnight. The dried tubes were moved to a desiccator for 2–3 days and then weighed on a microbalance. The dry cell weights were reported as g L^{-1} (g DCW per L of culture).

Residual glycerol analysis

Residual glycerol in the culture supernatants was analysed spectroscopically at 340 nm using a glycerol quantitation kit (r-biopharm, Roche) according to the manufacturer's protocol. Triplicate determinations were performed for each sample.

Residual glucose analysis

Residual glucose in the culture supernatants was analysed using an Accu-Chek Active Glucose analyser (Roche Diagnostics) or by Amplex® Red Glucose/Glucose Oxidase Assay Kit (Life Technologies) according to the manufacturer's protocol. Triplicate determinations were performed for each sample.

Residual methanol analysis

Residual methanol in the culture supernatants was analysed using a Thermo Scientific FOCUS Gas Chromatograph. Appropriately-diluted culture supernatants (1 μL) were injected in duplicate and the methanol peaks integrated using QuanLab® software. The mean value of the integrated peak area was used to estimate the residual methanol within the sample by comparison with methanol standards. Triplicate determinations were performed for each sample.

Preparations of membrane fractions

Harvested cells producing recombinant $hA_{2a}R$, hCD81, hCD82 or human claudin-1 were re-suspended in 2 mL breaking buffer (50 mM Na_2HPO_4, 2 mM EDTA, 100 mM NaCl, 5% glycerol w/v, pH 7.4) per gram of wet cell weight and disrupted at 30,000 psi for 20 min using an Avestin® C3 pressure homogeniser. Breaking efficiency was typically > 90% as determined by light microscopy. The samples were clarified by centrifugation at 10,000 × g, 4°C for 30 min and total membranes recovered from the supernatant at 100,000 × g, 4°C for 1 h. Total membranes were re-suspended in 10 mL Buffer A (20 mM HEPES, 50 mM NaCl, 10% glycerol, pH 7.0) per gram of total membrane. Protein concentration was determined using a BCA assay (Sigma-Aldrich) using BSA as a protein standard. 0.5 mL aliquots of membrane fraction were stored at −80°C prior to further analysis. Analysis by immunoblot was done as described previously [23].

Radio-ligand binding assays

Membrane bound $hA_{2a}R$ was analysed using a single-point saturation radio-ligand binding assay. Membranes

Motif Exhibit 1059, 426 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

(0.5 mg mL^{-1}) supplemented with 0.1 U of adenosine deaminase (Sigma-Aldrich) were incubated with 10 nM [^{3}H]-ZM241385 (American Radiolabelled Chemicals UK Ltd) for 1.5 h at 30°C. 10 nM of [^{3}H]-ZM241385 is expected to saturate the receptor [18]. To assay non-specific binding, 1 μM ZM241385 (Tocris) was added to the incubations. Assays were terminated by centrifugation at 14,000 rpm in a bench-top centrifuge for 5 min at 4°C. The supernatant was discarded and the pellets were washed with water and then solublilised with Soluene® (PerkinElmer) overnight. Solubilised pellets were added to scintillation fluid and counted in a scintillation counter (Packard 1600TR Liquid Scintillation Analyser) to determine the amount of bound radio-ligand. Triplicate determinations were performed for each sample. B_{max} values were calculated from the Langmuir occupancy equation and reported as pmol $hA_{2a}R$ per mg of total membrane protein [18], as determined by BCA assay.

Recombinant GFP analysis

Supernatants (100 μL) from cultures secreting GFP were assayed for GFP fluorescence using a Spectramax Gemini XS® plate reader with an excitation wavelength (λ_{exe}) of 397 nm and emission wavelength (λ_{em}) of 506 nm at 25°C. Triplicate determinations were performed for each sample. All samples and blanks were buffered to pH >7.0 using 50 μL 1 M potassium phosphate pH 8.0. To determine the concentration of GFP in each of the samples, a recombinant GFP standard (Vector Laboratories Ltd) was used to construct a standard curve relating GFP fluorescence to protein concentration, as previously described [22].

Recombinant CGRP-RCP-GFP analysis

Recombinant CGRP-RCP-GFP was analysed from culture supernatants by immunoblot as described previously [23].

Recombinant HRP analysis

Supernatants (25 μL) from cultures secreting HRP were assayed for HRP using a 2 Component ABTS Peroxidase Substrate system (Kirkegaard Perry Laboratories) containing ABTS [2,2′-azino-di-(3-ethylbenzthiazoline-6-sulfonate)] substrate and hydrogen peroxide. This liquid substrate system produces a blue-green coloured product when reacted with peroxidase present in the test samples. 25 μl of culture supernatant was mixed with 150 μl of ABTS-H_2O_2 in a 96 well plate and the increase in absorbance of the end product (after 10 minutes of incubation at room temperature) was read at 405 nm using a spectrophotometer (Multiskan GO, Thermo Scientific). To determine the concentration of HRP in each of the samples, a purified HRP standard (Sigma-Aldrich) was used to generate a standard curve relating peroxidase activity to protein concentration.

Graphical presentation

Graphs were constructed using OriginPro 9.1® software.

Additional file

Additional file 1: Supplementary data for bioreactor and shake-flask cultivations. Table S1. hCGRP-RCP-GFP is present in all phases of 1 L glycerol-grown *P. pastoris* bioreactor cultivations. **Table S2.** GFP is not produced in the pre-induction phases of glycerol- or glucose-grown *P. pastoris* shake-flask cultivations. **Table S3.** HRP is not produced in the pre-induction phases of glycerol-grown *P. pastoris* shake-flask cultivations. **Figure S1.** hCD81, hCD82 and human claudin-1 are not produced in the pre-induction phases of *P. pastoris* shake-flask cultivations as determined by immunoblot.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

ZB was involved in all aspects of the experimental design, data collection, data analysis and interpretation and was supported by RMB and MG. DRP supported the execution and analysis of the radio-ligand binding assays. SJR supported the execution and analysis of the hA2aR-, GFP- and hCGRP-RCP-GFP -producing cultivations. MJ, MC and DS produced data for HRP, hCD81, hCD82 and human claudin-1. ID provided hCGRP-RCP-GFP cDNA and the anti- hCGRP-RCP antibody. RMB directed the study, coordinated the data analysis and interpretation and, together with ZB, drafted the manuscript. All authors contributed to, read and approved the final version of the manuscript.

Acknowledgements

This work was supported by a BBSRC CASE studentship with AstraZeneca Ltd to RMB supporting ZB and by BBSRC grant BB/I019960/1 to DRP and RMB supporting SJR. The hCGRP-RCP cDNA was cloned and the antibody raised with support from NIH grant DK52328 to ID. We thank Jiteen Ahmed and Victoria Ewart for their invaluable help with the gas chromatography experiments and Aston Research Centre for Healthy Ageing (ARCHA) for support.

Author details

[1]School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham B4 7ET, United Kingdom. [2]AstraZeneca Ltd, Alderley Park, Macclesfield, Cheshire SK10 4TF, United Kingdom. [3]School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. [4]Department of Neurobiology and Anatomy, University of Rochester Medical Center, Rochester, New York 14642, USA. [5]present address for MG, Lonza Group Ltd, Muenchensteinerstrasse 38, CH-4002 Basel, Switzerland.

Received: 5 February 2014 Accepted: 18 August 2014
Published: 4 September 2014

References

1. De Schutter K, Lin Y, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouze P, Van de Peer Y, Callewaert N: **Genome sequence of the recombinant protein production host *Pichia pastoris*.** *Nat Biotechnol* 2009, **27:**561–566.
2. Mattanovich D, Graf A, Stadlmann J, Dragosits M, Redl A, Maurer M, Kleinheinz M, Sauer M, Altmann F, Gasser B: **Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*.** *Microb Cell Fact* 2009, **8:**29–29.
3. Lin-Cereghino J, Cregg J: **Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*.** *FEMS Microbiol Rev* 2000, **24:**45–66.
4. Cos O, Ramón R, Montesinos J, Valero F: **Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast Pichia pastoris under different promoters: A review.** *Microb Cell Fact* 2006, **5:**17.
5. Cereghino G, Cereghino J, Ilgen C, JM C: **Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*.** *Curr Opin Biotechnol* 2002, **13:**329–332.
6. Minning S, Serrano A, Ferrer P, Sola C, Schmidt R, Valero F: **Optimization of the high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris*.** *J Biotechnol* 2001, **86:**59–70.

Motif Exhibit 1059, 427 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

7. Filmore D: **It's a GPCR world.** *Modern Drug Discov* 2004, **7:**24–28.
8. Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, Abagyan R, Cherezov V, Liu W, Han G, Kobayashi T, Stevens RC, Iwata S: **Structure of the human histamine H1 receptor complex with doxepin.** *Nature* 2011, **475:**65–72.
9. Hino T, Arakawa T, Iwanari H, Yurugi-Kobayashi T, Ikeda-Suno C, Nakada-Nakura Y, Kusano-Arai O, Weyand S, Shimamura T, Nomura N, Cameron AD, Kobayashi T, Hamakubo T, Iwata S, Murata T: **G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody.** *Nature* 2012, **482:**237–240.
10. Schmidt F: **Optimization and scale up of industrial fermentation processes.** *Appl Microbiol Biotechnol* 2005, **68:**425–435.
11. André N, Cherouati N, Prual C, Steffan T, Zeder-Lutz G, Magnin T, Pattus F, Michel H, Wagner R, Reinhart C: **Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen.** *Protein Sci* 2006, **15:**1115–1126.
12. Lundstrom K, Wagner R, Reinhart C, Desmyter A, Cherouati N, Magnin T, Zeder-Lutz G, Courtot M, Prual C, André N, Hassaine G, Michel H, Cambillau C, Pattus F: **Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems.** *J Struct Funct Genom* 2006, **7:**77–91.
13. Singh S, Hedley D, Kara E, Gras A, Iwata S, Ruprecht J, Strange P, Byrne B: **A purified C-terminally truncated human adenosine A(2A) receptor construct is functionally stable and degradation resistant.** *Prot Expr Purif* 2010, **74:**80–87.
14. Singh S, Gras A, Fiez-Vandal C, Ruprecht J, Rana R, Martinez M, Strange P, Wagner R, Byrne B: **Large-scale functional expression of WT and truncated human adenosine A2A receptor in *Pichia pastoris* bioreactor cultures.** *Microb Cell Fact* 2008, **7:**1–10.
15. Singh S, Gras A, Fiez-Vandal C, Martinez M, Wagner R, Byrne B: **Screening for high-yielding Pichia pastoris clones: the production of G protein-coupled receptors as a case study.** *Methods Mol Biol* 2012, **866:**65–73.
16. Inan M, Meagher M: **The effect of ethanol and acetate on protein expression.** *Biotechnol Bioeng* 2001, **92:**337–341.
17. Hellwig S, Emde F, Raven N, Henke M, van der Logt P, Fischer R: **Analysis of single-chain antibody production in *Pichia pastoris* using on-line methanol control in fed-batch and mixed-feed fermentations.** *Biotechnol Bioeng* 2001, **74:**344–352.
18. Fraser NJ: **Expression and functional purification of a glycosylation deficient version of the human adenosine 2a receptor for structural studies.** *Protein Expr Purif* 2006, **49:**129–137.
19. Weinhandl K, Winkler M, Glieder A, Camattari A: **Carbon source dependent promoters in yeasts.** *Microb Cell Fact* 2014, **13:**5.
20. Potvin G, Ahmad A, Zhang Z: **Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: A review.** *Biochem Eng J* 2012, **64:**91–105.
21. Luebke AE, Dahl GP, Roos BA, Dickerson IM: **Identification of a protein that confers calcitonin gene-related peptide responsiveness to oocytes by using a cystic fibrosis transmembrane conductance regulator assay.** *Proc Natl Acad Sci U S A* 1996, **93:**3455–3460.
22. Holmes W, Darby R, Wilks M, Smith R, Bill R: **Developing a scalable model of recombinant protein yield from *Pichia pastoris*: the influence of culture conditions, biomass and induction regime.** *Microb Cell Fact* 2009, **8:**35.
23. Bonander N, Jamshad M, Oberthur D, Clare M, Barwell J, Hu K, Farquhar MJ, Stamataki Z, Harris HJ, Dierks K, Dafforn TR, Betzel C, McKeating JA, Bill RM: **Production, purification and characterization of recombinant, full-length human claudin-1.** *PLoS One* 2013, **8:**e64517.
24. Çelik E, Çalık P, Oliver S: **Fed-batch methanol feeding strategy for recombinant protein production by *Pichia pastoris* in the presence of co-substrate sorbitol.** *Yeast* 2009, **26:**473–484.
25. Çelik E, Çalık P, Oliver S: **Metabolic flux analysis for recombinant protein production by *Pichia pastoris* using dual carbon sources: effects of methanol feeding rate.** *Biotechnol Bioeng* 2010, **105:**317–329.
26. Jungo C, Schenk J, Pasquier M, Marison I, von Stockar U: **A quantitative analysis of the benefits of mixed feeds of sorbitol and methanol for the production of recombinant avidin with *Pichia pastoris*.** *J Biotechnol* 2007, **131:**57–66.
27. Kobayashi K, Kuwae S, Ohya T, Ohda T, Ohyama M, Tomomitsu K: **High level secretion of recombinant human serum albumin by fedbatch fermentation of themethylotrophic yeast, *Pichia pastoris*, based on optimal methanol feeding strategy.** *J Biosci Bioeng* 2000, **90:**280–288.
28. Zhang W, Sinha J, Smith L, Inan M, Meagher M: **Maximization of production of secreted recombinant proteins in *Pichia pastoris* fed-batch fermentation.** *Biotechnol Progr* 2005, **21:**386–393.
29. Routledge SJ, Hewitt CJ, Bora N, Bill RM: **Antifoam addition to shake flask cultures of recombinant Pichia pastoris increases yield.** *Microb Cell Fact* 2011, **10:**17.
30. Cunha A, Clemente J, Gomes R, Pinto F, Thomaz M, Miranda S, Pinto R, Moosmayer D, Donner P, Carrando M: **Methanol induction optimization for scFv antibody fragment production in *Pichia pastoris*.** *Biotechnol Bioeng* 2004, **86:**458–467.
31. Schenk J, Marison I, von Stockar U: **A simple method to monitor and control methanol feeding of *Pichia pastoris* fermentations using mid-IR spectroscopy.** *J Biotechnol* 2007, **128:**344–353.
32. Çalık P, İnankur B, Soyaslan E, Şahin M, Taşpınar H, Açık E, Bayraktar E: **Fermentation and oxygen transfer characteristics in recombinant human growth hormone production by *Pichia pastoris* in sorbitol batch and methanol fed-batch operation.** *J Chem Technol Biotech* 2010, **85:**226–233.
33. Zhang W, Bevins M, Plantz BA, Smith L, Meagher M: **Modeling *Pichia pastoris* growth on methanol and optimizing the production of a recombinant protein, the heavy-chain fragment C of botulinum neurotoxin, serotype A.** *Biotechnol Bioeng* 2000, **70:**1–8.
34. Inan M, Meagher MM: **Non-repressing carbon sources for alcohol oxidase (AOX1) promoter of Pichia pastoris.** *J Biosci Bioeng* 2001, **92:**585–589.
35. Ferndahl C, Bonander N, Logez C, Wagner R, Gustafsson L, Larsson C, Hedfalk K, Darby RA, Bill RM: **Increasing cell biomass in Saccharomyces cerevisiae increases recombinant protein yield: the use of a respiratory strain as a microbial cell factory.** *Microb Cell Fact* 2010, **9:**47.

doi:10.1186/s12934-014-0127-y
Cite this article as: Bawa *et al.*: Functional recombinant protein is present in the pre-induction phases of *Pichia pastoris* cultures when grown in bioreactors, but not shake-flasks. *Microbial Cell Factories* 2014 **13**:127.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central

Ex. 1037 Ferreira 2013

ENCYCLOPEDIA OF BIOLOGICAL CHEMISTRY

EDITORS-IN-CHIEF

WILLIAM J. LENNARZ

Stony Brook University,
Stony Brook, NY, USA

M. DANIEL LANE

Johns Hopkins University School of Medicine,
Baltimore, MD, USA

VOLUME 2

DNA–L



ELSEVIER

AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



ACADEMIC PRESS

Academic Press is an imprint of Elsevier
32 Jamestown Road, London NW1 7BY, UK
30 Corporate Drive, Suite 400, Burlington MA 01803, USA
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA

First edition 2004
Second edition 2013

Copyright © 2013 Elsevier, Inc. All rights reserved.

The following articles are US Government works in the public domain and are not subject to copyright:

Adipocyte Fat Mobilization: The Regulatory Roles of Perilipin 1, Adipose Triglyceride Lipase, and Hormone-Sensitive Lipase
Arf Family
Calcium Signaling: Calmodulin-Dependent Phosphatase
Cyclic AMP Receptors of *Dictyostelium*
Cyclic Nucleotide Phosphodiesterases
Cytokines
DNA Mismatch Repair in Disease and Aging
DNA Polymerase β, Eukaryotic
Endocytosis
Eukaryotic Chemotaxis
Gs Family of Heterotrimeric G Proteins
Inositol Phosphate Kinases and Phosphatases
Phospholipase C
Prions of Yeast and Fungi: Proteins Acting as Genes
Riboswitches
Structure and Regulation of Pyruvate Dehydrogenases

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher.

Permissions may be sought from Elsevier's Science & Technology Rights department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com.

Alternatively you can submit your request online by visiting the Elsevier website at http://elsevier.com/locate/permissions and selecting *Obtaining permission to use Elsevier material.*

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-378630-2

For information on all Elsevier publications
visit our website at store.elsevier.com

Printed and bound in the United States of America

13 14 15 16 17 10 9 8 7 6 5 4 3 2

20134319 94

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Editorial: Jo Collett, Cari Owen
Production: Justin Taylor

CONTENTS

Editors-in-Chief *v*

Editorial Board *vii*

Contributors *ix*

How to Use the Encyclopedia *xxxiii*

Preface *xxxv*

VOLUME 1

AAA-ATPases 1
J Martin and A N Lupas

ABC Transporters 7
A Goffeau and B De Hertogh

Actin Assembly/Disassembly 12
H N Higgs

Actin-Capping and -Severing Proteins 18
S Maiti and J R Bamburg

Actin Organization 27
T Svitkina

Actin-Related Proteins 36
R D Mullins

Adenylyl Cyclases 42
R Taussig

Adipocyte Fat Mobilization: The Regulatory Roles of Perilipin 1, Adipose Triglyceride Lipase, and Hormone-Sensitive Lipase 47
K T Dalen, C Londos, and A R Kimmel

Adipogenesis 52
M D Lane

Adrenergic Receptors 57
D B Bylund

Affinity Tags for Protein Purification 61
J J Falke and J A Corbin

Algal Hydrogen Production 66
U M N Murthy and M L Ghirardi

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Allosteric Regulation 71
B S Cooperman

Alternative Splicing 75
K Lynch and L Motta-Mena

Alternative Splicing: Regulation of *Drosophila melanogaster* Sex Determination 81
J K M Penn, P Graham, P Schedl, and H K Salz

Amine Oxidases 87
G Floris and A Finazzi Agrò

Amino Acid Metabolism 91
L Cynober

Aminopeptidases 97
R A Bradshaw

Amyloid 100
R Wetzel

Anaplerosis 105
R R Russell III and H Taegtmeyer

Angiotensin Receptors 111
T Inagami

Anorexigenic and Orexigenic Gut Peptides 115
P S Petersen and T H Moran

Aquaporin Water Channels 120
K Liu and P Agre

Arachidonic Acid Regulated Calcium Channel 127
T J Shuttleworth

Arf Family 134
G Pacheco-Rodriguez, J Moss, and M Vaughan

Aspartic Proteases 137
B M Dunn

ATP in Plant Mitochondria: Substrates, Inhibitors, and Uncouplers 141
K L Soole and R I Menz

ATP Synthesis: Mitochondrial Cyanide-Resistant Terminal Oxidases 145
J N Siedow

Autophagy in Fungi and Mammals 149
J Geng, D J Klionsky, and J-L Guan

B_{12}-Containing Enzymes 155
V Bandarian and R G Matthews

B-Cell Antigen Receptor 161
T M Yankee

Bcl2 Family: Cell Death Enhancers and Inhibitors 165
D L Vaux

Bile Salts 167
L Maillette de Buy Wenniger, T Pusl, and U Beuers

Biochemistry: Niacin/NAD(P) 172
K L Bogan and C Brenner

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Biochemistry of Development: Striated Muscle 179
R Bassel-Duby and E N Olson

Biochemistry of Hematopoiesis 187
A D Friedman

Biochemistry of Liver Regeneration 190
B Ochoa and A M Diehl

Biochemistry of Neurogenesis 196
K Christian, H Song, and G-L Ming

Biochemistry of Thiamine and Thiamine Phosphate Compounds 202
L Bettendorff and P Wins

Bioenergetics and the Mitochondrial Genome 210
D C Wallace

Bioenergetics: General Definition of Principles 217
D G Nicholls

Biotin 221
S W Polyak and A Chapman-Smith

Biotinylation of Proteins 226
R A Kohanski

Bmp Signaling and Vascular Disease 229
J Lowery and M de Caestecker

Bradykinin Receptors 240
R M Burch

Branched-Chain Amino Acids 244
D T Chuang

Brassinosteroids 250
S D Clouse

Cadherin-Mediated Cell–Cell Adhesion 255
M Lowndes and W J Nelson

Cadherin Signaling 261
D B Sacks

Calcitonin Gene-Related Peptide and Adrenomedullin Receptors 265
D L Hay, A C Conner, and D R Poyner

Calcitonin Receptor 270
S I Girgis, N Moradi-Bidhendi, and L Mancini

Calcium-Binding Proteins: Cytosolic (Annexins, Gelsolins, and C_2-Domain Proteins) 275
J Krebs

Calcium, Biological Fitness of 281
R J P Williams

Calcium-Buffering Proteins: Calbindin 284
L J Mady, F Haleem, and S Christakos

Calcium-Buffering Proteins: ER Luminal Proteins 290
J Groenendyk and M Michalak

Calcium/Calmodulin-Dependent Protein Kinase II 297
A Hudmon and H Schulman

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Calcium/Calmodulin-Dependent Protein Kinases 304
A J Robison and R J Colbran

Calcium in the Regulation of Gene Expression 310
M Rivas and J R Naranjo

Calcium-Modulated Proteins (EF-Hand) 316
R H Kretsinger

Calcium Oscillations 322
M Brini

Calcium-Sensing Receptor (CaSR) 326
J Tfelt-Hansen and E M Brown

Calcium Signaling by Cyclic ADP-Ribose and NAADP 331
L Santella and J T Chun

Calcium Signaling: Calmodulin-Dependent Phosphatase 337
C B Klee and S Li

Calcium Signaling: NO Synthase 342
Z-Q Wang and D J Stuehr

Calcium Transport in Mitochondria 347
R Rizzuto and M Brini

Calpain 353
H Sorimachi, S Hata, and Y Ono

Carbohydrate Chains: Enzymatic and Chemical Synthesis 362
T J Tolbert and C-H Wong

Carbohydrate Metabolism in the Central Nervous System 368
I Simpson and S J Vannucci

Carbohydrate-Responsive Element Binding Protein 374
K Uyeda and H Yamashita

Cardiac Excitation–Contraction Coupling 379
D M Bers and S Despa

Carnitine and β-Oxidation 384
J Kerner and C L Hoppel

Caspases and Cell Death 388
D W Nicholson and G Melino

Catalysis by RNA, Structural Themes: Group I Introns 397
B L Golden

Cell Cycle: Control of Entry and Progression through S Phase 401
S L Forsburg

Cell-Cycle Controls in G_1 and G_0 406
A Narasimha, W Shi, and S F Dowdy

Cell Cycle: DNA Damage Checkpoints 410
M D Weitzman and J Y J Wang

Cell Cycle: Mitotic Checkpoint 417
T J Yen

Cell Death by Apoptosis and Necrosis 423
V Ermer and P Nicotera

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Cell–Matrix Interactions 426
J A Askari and M J Humphries

Cell Migration 430
J V Small

Cell Migration within Three-Dimensional Matrices 436
P J Keely

Centromeres 446
B A Sullivan

c-*fes* Proto-Oncogene 451
T E Smithgall

Chaperonins 456
A Bracher and F U Hartl

Chemiluminescence and Bioluminescence 461
T O Baldwin

Chemiosmotic Theory 467
P R Rich

Chemistry of Alzheimer Disease 473
G H Sack Jr.

Chemokine Receptors 480
T Sobolik-Delmaire, D Raman, J Sai, G-H Fan, and A Richmond

Chemolithotrophy 486
A B Hooper and A A DiSpirito

Chemotactic Peptide/Complement Receptors 493
E R Prossnitz and L A Sklar

Chlorophylls and Carotenoids 498
H Scheer

Chloroplasts 506
N Rascio

Chloroplast Signaling and Redox Control 511
J-D Rochaix

Cholesterol Synthesis and Regulation 516
P J Espenshade

Chromatin: Methyl-CpG-Binding Proteins 521
D G Skalnik

Chromatin: Nucleosome Positioning – the GAL Promoter 525
D Lohr

Chromatin: Physical Organization 530
C L Woodcock

Chromosome Organization and Structure, Overview 535
E Gracheva and S C R Elgin

Coenzyme A 539
M D Lane

Collagenases 542
K Holmbeck and H Birkedal-Hansen

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Collagens 545
D J Prockop

Color Vision 550
G H Jacobs

Conservative Site-Specific Recombination 555
M C M Smith

Cyclic AMP Receptors of *Dictyostelium* 562
D Hereld, Y Kamimura, and P N Devreotes

Cyclic GMP Phosphodiesterases 567
S H Francis and J D Corbin

Cyclic Nucleotide-Dependent Protein Kinases 574
S H Francis and J D Corbin

Cyclic Nucleotide Phosphodiesterases 580
V C Manganiello and E Degerman

Cyclic Nucleotide-Regulated Cation Channels 586
M Biel

Cysteine Proteases 589
D J Buttle and J S Mort

Cytochrome bc_1 Complex (Respiratory Chain Complex III) 593
B L Trumpower

Cytochrome *c* 599
H Tuppy and G Kreil

Cytochrome Oxidases, Bacterial 602
P Brzezinski and P Ädelroth

Cytochrome P-450 607
R Bernhardt

Cytokines 613
S J Levine

Cytokinesis 622
V Srivastava and D N Robinson

Cytokinin 627
T Schmülling

Cytoskeletal Motors: General Principles 632
R S Rock

Desmosomes and Hemidesmosomes 636
V Todorović, K R Kligys, R L Dusek, J C R Jones, and K J Green

Detergent Properties 644
D R McCaslin

Diabetes 649
S M Blackman and D W Cooke

Diacylglycerol Kinases and Phosphatidic Acid Phosphatases 659
M K Topham

Disulfide Bond Formation 664
H F Gilbert

Motif Exhibit 1059, 437 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

VOLUME 2

DNA Base Excision Repair Pathways 1
R A O Bennett and B Demple

DNA Damage: Alkylation 9
J L Tubbs and J A Tainer

DNA Glycosylases: Mechanisms 16
A C Drohat and A Maiti

DNA Helicases: Hexameric Enzyme Action 21
D Sen and S S Patel

DNA Ligases: Mechanism and Functions 28
A E Tomkinson and J A Della-Maria

DNA Ligases: Structures 33
J M Pascal

DNA Methyltransferases: Eubacterial GATC 40
M G Marinus

DNA Mismatch Repair and Homologous Recombination 45
I Matic

DNA Mismatch Repair and the DNA Damage Response 51
G-M Li

DNA Mismatch Repair in Bacteria 54
A-L Lu

DNA Mismatch Repair in Disease and Aging 60
P Hsieh

DNA Mismatch Repair in Mammals 66
A J Plys and E Alani

DNA Modification, Restriction Endonucleases: Type I Enzymes 72
D T F Dryden and G A Roberts

DNA Oxidation 77
D O Zharkov

DNA Polymerase β, Eukaryotic 82
W A Beard and S H Wilson

DNA Polymerase I, Bacterial 87
C M Joyce

DNA Polymerase III, Bacterial 92
H Maki and A Furukohri

DNA Polymerases: Kinetics and Mechanism 96
K A Johnson

DNA Polymerases: Reverse Transcriptase Integrase, and Retrovirus Replication 101
M-L Andréola, V Parissi, and S Litvak

DNA Replication: Eukaryotic Origins and the Origin Recognition Complex 108
I Chesnokov and A Svitin

DNA Replication Fork, Bacterial 114
M M Spiering and S J Benkovic

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

DNA Replication Fork, Eukaryotic 118
L M Kelman and Z Kelman

DNA Replication: Initiation in Bacteria 121
J M Kaguni

DNA Replication, Mitochondrial 126
D A Clayton

DNA Restriction and Modification: Type II Enzymes 130
S E Halford

DNA Restriction and Modification: Type III Enzymes 136
D N Rao and S Bheemanaik

DNA Secondary Structure 142
A Bacolla, K M Vasquez, and R D Wells

DNA Sequence Recognition by Proteins 149
K Rutherford and G D Van Duyne

DNA Supercoiling 154
T-S Hsieh

DNA Topoisomerases: Type I 157
J J Champoux

DNA Topoisomerases: Type II 163
A C Gentry and N Osheroff

Dopamine Receptors 169
K A Neve

Dynactin 174
T A Schroer

Dynamic Behavior of Photosystem II Light Harvesting 178
P Horton

Dynein 184
K K Pfister

Eicosanoid Receptors 188
R M Breyer and M D Breyer

Elastin 192
J A Foster

Endocannabinoids 194
D Piomelli

Endocytosis 197
J G Donaldson

Endoplasmic Reticulum-Associated Protein Degradation 200
R Bernasconi and M Molinari

Energy Transduction in Anaerobic Bacteria 204
G Unden

Enzyme Inhibitors 210
V L Schramm

Enzyme Kinetics 216
I H Segel

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Enzyme Reaction Mechanisms: Stereochemistry 221
M-D Tsai, L Zhao, and B J Lamarche

ER/SR Calcium Pump: Function 226
G Inesi

ER/SR Calcium Pump: Structure 231
C Toyoshima

Eukaryotic Chemotaxis 236
C P McCann and C A Parent

Eukaryotic DNA Polymerase α 241
C Lee and B K Tye

Eukaryotic DNA Polymerase δ 245
P M J Burgers

Eukaryotic Protein Biosynthesis: The Elongation Cycle 249
A A Komar, D J Taylor, and W C Merrick

Evolving Concepts of Leptin 256
H Münzberg and T W Gettys

Exosomes and Microvesicles 262
S J Gould

FAK Family 265
S K Hanks

F-ATPases 269
J E Walker

Fatty Acid Metabolism and Cancer 275
F P Kuhajda

Fatty Acid Oxidation 281
H Schulz

Fatty Acid Structure and Synthesis 285
S D Clarke and M T Nakamura

Fatty Acyl-CoA Synthetases 290
P A Watkins

Ferredoxin 296
G Zanetti and V Pandini

Ferredoxin-NADP$^+$ Reductase 299
G Zanetti and A Aliverti

Fibroblast Growth Factor Receptors and Cancer-Associated Perturbations 303
M Kornmann and M Korc

Flavins 309
B Entsch and D P Ballou

Flippases 314
C J Waechter

Focal Adhesions and Related Integrin Contacts 318
E Zamir and B Geiger

Folate and Vitamin B_{12} Function 324
B Shane

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Free Oligosaccharides: Formation and Degradation (Free Oligosaccharides Related to N-Linked Glycans) 329
T Suzuki

Friedreich's Ataxia 332
P E Hart and A H V Schapira

F_X, F_A, and F_B Iron–Sulfur Clusters in Type I Photosynthetic Reaction Centers 335
B Jagannathan and J H Golbeck

G_{12}/G_{13} Family 343
S Offermanns

$GABA_A$ Receptor 347
R W Olsen and G-D Li

$GABA_B$ Receptor 351
S J Enna and N G Bowery

Gastrointestinal Digestion and Absorption 354
J Keller

Gene Expression in Bacterial Systems: The *trp* Operon and Attenuation 360
P Gollnick

Gene Expression in Eukaryotes: RNA Polymerase II Structure 365
P Cramer

Genome-Wide Analysis of Gene Expression 369
D-W Doug Chung and K G Le Roch

Giant Mitochondria (Megamitochondria) 375
B Tandler and C L Hoppel

Gi Family of Heterotrimeric G Proteins 377
M E Linder

Gluconeogenesis 381
R W Hanson and O E Owen

Glucose/Sugar Transport in Bacteria 387
L Guan and H R Kaback

Glucose/Sugar Transport in Mammals 391
S Mora and J Pessin

Glutamate Receptors, Metabotropic 395
K J Gregory and P J Conn

Glutathione Peroxidases 399
F Ursini and M Maiorino

Glycation 405
M Lima and J W Baynes

Glycerolipid Structure, Function, and Synthesis in Eukaryotes 412
D R Voelker

Glycine Receptors 419
T Lynagh and B Laube

Glycogen Metabolism 425
P J Roach

Glycogen Storage Diseases 428
G H Sack Jr.

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Glycogen Synthase Kinase-3 434
J R Woodgett

Glycolipid-Dependent Adhesion Processes 439
S Hakomori

Glycolysis Overview 443
R A Harris

Glycoprotein Folding and Processing Reactions 448
J J Caramelo and A J Parodi

Glycoprotein-Mediated Cell Interactions, Nonmucin-Type O-Linked 453
R S Haltiwanger

Glycoproteins, N-Linked 457
M A Lehrman

Glycoproteins, Plant 465
C J Schultz

Glycosylation, Congenital Disorders of 469
B G Ng and H H Freeze

Glycosylphosphatidylinositol Anchors 476
A K Menon

Golgi Complex 479
M Stamnes

G-Protein-Coupled Receptor Kinases and Arrestins 483
N Zhukovskaya, B Barker, D S Kang, and J L Benovic

G Protein Signaling Regulators 492
J H Exton

G_q Family 496
W Yang and J D Hildebrandt

Green Bacteria: Chlorophyll Biosynthesis, Light-Harvesting, Reaction Centers, and Electron Transport 501
D A Bryant

Green Bacteria: Secondary Electron Donor (Cytochromes) 510
H Oh-oka and R E Blankenship

Green Bacteria: The Light-Harvesting Chlorosome 513
J M Olson

Gs Family of Heterotrimeric G Proteins 517
L S Weinstein

Heat/Stress Responses 522
G Thibault and D T W Ng

Hematopoietin Receptors 526
B A Miller

Heme Proteins 532
J Everse

Heme Synthesis 539
G C Ferreira

Hexokinase/Glucokinase 543
E Van Schaftingen

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

HIV-1 Reverse Transcriptase Structures 548
K Das, E Arnold, and S H Hughes

HIV Protease 554
B M Dunn

Homologous Recombination in Meiosis 557
N M Hollingsworth

Hydrogenases, Structure and Function 562
W Lubitz and H Ogata

Immunoglobulin (Fc) Receptors 568
M Hogarth

Indicators of Intracellular Calcium 574
M Giacomello and M Campeol

Inositol Phosphate Kinases and Phosphatases 580
S B Shears

Inositol Trisphosphate and Calcium Signaling 583
M J Berridge

Insect Biochemistry/Hormones 590
P Y Scaraffia and R L Miesfeld

Insulin- and Glucagon-Secreting Cells of the Pancreas 596
F M Matschinsky

Insulin Mechanisms/Metabolic Actions 602
E J Gallagher and D LeRoith

Integrin Signaling 608
L E Goldfinger

Intercellular Ca^{2+} Waves: Mechanisms of Initiation and Propagation 613
L Leybaert and M J Sanderson

Interferon Receptors 619
C P Elco and G C Sen

Intermediate Filament Linker Proteins: Plectin and BPAG1 624
P Fuchs and G Wiche

Intermediate Filaments 631
K M Bernot and P A Coulombe

Intracellular Calcium Channels: $NAADP^{+}$-Modulated 637
A A Genazzani and M Debidda

Intracellular Calcium Waves 640
J T Chun and L Santella

Ion Channel Protein Superfamily 648
W A Catterall

IP_3 Receptors 653
K Mikoshiba and C Hisatsune

Iron–Sulfur Proteins 657
R Cammack

Keratins and the Skin 665
P A Coulombe, K M Bernot, and C-H Lee

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Kinesins as Microtubule Disassembly Enzymes 672
R Ohi and L Wordeman

Kinesin Superfamily Proteins 679
N Hirokawa and R Takemura

Kinetic Isotope Effects 688
J P Roth and J P Klinman

lac Operon 694
L Swint-Kruse and K S Matthews

Lectins 701
N Sharon and H Lis

LexA Regulatory System 706
M Butala, D Žgur-Bertok, and S J W Busby

Ligand-Operated Membrane Channels: Calcium (Glutamate) 712
E K Michaelis

Ligand-Operated Membrane Channels: GABA 722
F Minier and E Sigel

Light-Harvesting Complex I and II: Pigments and Proteins 726
S Jansson

Lipases 729
H L Brockman

Lipid Bilayer Structure 733
E London

Lipid Modification of Proteins: Targeting to Membranes 736
M D Resh

Lipid Rafts 741
D A Brown

Lipid Signaling and Ion Channels 745
J B Jensen, B H Falkenburger, and B Hille

Live Cell Imaging of Nuclear Dynamics 749
J C Harr and K L Reddy

Low-Barrier Hydrogen Bonds 756
P A Frey

Luft's Disease 760
S DiMauro

Lysophospholipid Receptors 762
G J Tigyi

VOLUME 3

Major Sperm Protein and Sperm Locomotion 1
K Shimabukuro and T M Roberts

Mammalian DNA Methyltransferase Structural Themes 6
X Cheng and H Hashimoto

Mass Spectrometry of Native Complexes 11
R J Rose and A J R Heck

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Meiosis 17
N Hunter

Melanocortin System 24
J Pantel and R D Cone

Membrane-Associated Energy Transduction in Bacteria and Archaea 28
G Schäfer

Membrane Fusion 36
R K Plemper and G B Melikyan

Membrane Transporters: Na^+/Ca^{2+} Exchangers 44
J Lytton

Membrane Transport, General Concepts 49
S G Schultz

Messenger RNA Degradation in Bacteria 52
D H Bechhofer

Messenger RNA Processing in Eukaryotes 59
K Potter, N Cremona, and J A Wise

Metabolic Control during Ischemia of the Heart 65
J R Ussher and G D Lopaschuk

Metabolic Roles of Adiponectin 72
K Sun and P E Scherer

Metabolic Roles of the Orexigenic and Anorexigenic Neuropeptides 76
G J Morton and M W Schwartz

Metabolite Channeling: Creatine Kinase Microcompartments 80
U Schlattner, M Tokarska-Schlattner, and T Wallimann

Metalloproteases 86
D S Auld

Metalloproteinases, Matrix 90
H Nagase and G Murphy

MicroRNAs in Eukaryotes 98
R N Douglas and J A Birchler

Microtubule-Associated Proteins 103
N Hirokawa and R Takemura

Mitochondria in Myocardial Ischemia 108
F Di Lisa

Mitochondrial Auto-Antibodies 113
H Baum

Mitochondrial Calcium Transport: Historical Aspects 118
E Carafoli

Mitochondrial Channels 127
M C Sorgato and A Bertoli

Mitochondrial DNA 132
G Schatz

Mitochondrial Dynamics 135
L Scorrano

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Mitochondrial Genes and Their Expression: Yeast 140
P P Slonimski and G Carignani

Mitochondrial Membranes, Structural Organization 145
C A Mannella

Mitochondrial Metabolite Carrier Protein Family 149
F Palmieri

Mitochondrial Outer Membrane and the VDAC Channel 159
M Colombini

Mitochondrial Permeability Transition Pore 162
P Bernardi

Mitochondria: Source and Target of Free Radicals 168
A Boveris and E Cadenas

Mitogen-Activated Protein Kinase Family 176
H Teramoto and J S Gutkind

Mitosis 181
P Wadsworth and J Titus

Monoamine Oxidase 185
A Mattevi

mRNA Polyadenylation in Eukaryotes 188
J L Manley and D C Di Giammartino

mTOR and its Downstream Targets 194
C G Proud

Mucin Family of Glycoproteins 200
R B Brown and M A Hollingsworth

Mucins in Embryo Implantation 205
N Dharmaraj and D D Carson

Multidrug Resistance Membrane Proteins 210
N C Rockwell

Muscarinic Acetylcholine Receptors 215
N M Nathanson

Myosin Motors 218
M A Titus and R E Larson

Natriuretic Peptides, Their Receptors and Therapeutic Applications 225
K Barbieri and L R Potter

Neoglycoproteins 229
Y C Lee and R T Lee

Neuronal Intermediate Filaments 233
R K H Liem

Neurotransmitter Transporters 238
K Erreger, H J G Matthies, A Galli, and C Saunders

Neurotrophin Receptor Signaling 241
R S Kenchappa and B D Carter

Nicotinamide Nucleotide Transhydrogenase 247
J Rydström

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Nitric Oxide Signaling 253
M A Marletta

N-Linked Glycan-Processing Glucosidases and Mannosidases 257
T Suzuki

Nonhexameric SF1 DNA Helicases/Translocases 261
E Antony and T M Lohman

Nonhomologous End Joining in Bacteria 266
E J Bartlett and A J Doherty

Nonhomologous End Joining in Eukaryotes 269
E Weterings and D J Chen

Nonhomologous Recombination: Bacterial Transposons 275
B T Hoang and M G Chandler

Nonhomologous Recombination: Retrotransposons 283
X Qi and S Sandmeyer

Nuclear Calcium in Synapse-to-Nucleus Communication: Common Regulator of Persistent Adaptations in the Nervous System 292
H Bading

Nuclear Compartmentalization 296
K P Smith and J B Lawrence

Nuclear Factor Kappa B 302
T D Gilmore

Nuclear Genes Involved in Mitochondrial Function and Biogenesis 306
A Tzagoloff and C L Dieckmann

Nuclear Lamina 310
B M Paschal and J B Kelley

Nuclear Organization, Chromatin Structure, and Gene Silencing 314
M W Vitalini, G Dialynas, L L Wallrath, and S R Mackey

Nuclear Pores and Nuclear Import/Export 318
A Lange and A H Corbett

Nucleolus, Overview 324
T Pederson

Nucleotide Excision Repair, Bacterial: The UvrABCD System 328
B Van Houten

Nucleotide Excision Repair: Biology 337
E C Friedberg

Nucleotide Excision Repair in Eukaryotes 341
N Goosen

Olfactory Receptors 345
S I Korsching

O-Linked GlcNAc Biosynthesis and Function 350
K Sakabe and G W Hart

Opioid Receptors 354
P Y Law and H H Loh

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

ORAI Store-Operated Calcium Channel 359
Y Zhou, P G Hogan, and A Rao

Organization of the Bacterial Nucleoid 367
C J Dorman

Oxygenases 371
O Hayaishi

P2Y Receptors 375
G R Dubyak

p53 Family 379
C E Barton and J A Pietenpol

Parathyroid Hormone/Parathyroid Hormone-Related Protein Receptor 384
T J Gardella

Pentose Phosphate (Hexose Mono Phosphate) Pathway 391
B R Landau

Peptide Amidation 396
M J Niciu, R E Mains, and B A Eipper

Periplasmic Electron-Transport Systems in Bacteria 400
D J Richardson, G Sawers, and R J M Van Spanning

Peroxidase Biochemistry and Redox Signaling 407
A Bindoli and M P Rigobello

Peroxisomal Metabolism 413
S J Gould

Peroxisome Proliferator-Activated Receptors 418
M C Sugden, P W Caton, and M J Holness

Peroxisomes 425
S Joshi and S Subramani

Phage Display for Protein Binding 431
H B Lowman

Phagocytosis 437
G Wright, K Srinivasan, and C Janetopoulos

Pheromone Receptors (Yeast) 441
J B Konopka and J W Thorner

Phosphatidylinositol-3-Phosphate 447
J V Virbasius and M P Czech

Phosphatidylinositol Bisphosphate and Trisphosphate 451
A Toker

Phosphofructokinase-2/Fructose Bisphosphatase-2 456
D M Raben

Phosphoinositide 4- and 5-Kinases and Phosphatases 459
S F Bairstow, M W Bunce, and R A Anderson

Phosphoinositide-Dependent Protein Kinases 463
N R Leslie and C P Downes

Phospholipase C 467
F Sekiya

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Phospholipase D 472
W Ali and M A Frohman

Phospholipid Synthesis in Yeast 478
G-S Han and G M Carman

Photoinhibition and Photoprotection in Plants, Algae, and Cyanobacteria 482
G M Giacometti and T Morosinotto

Photoreceptors 488
K-W Yau

Photosynthesis 492
R C Leegood

Photosynthetic Carbon Dioxide Fixation 497
M J Paul

Photosystem I 503
I Grotjohann and P Fromme

Photosystem II: Assembly and Turnover of the D1 Protein 508
S Sirpiö, E Kanervo, and E-M Aro

Photosystem II: Redox and Protein Components 512
J Barber

Photosystem II: Water Oxidation, Overview 520
F Rappaport and A Boussac

Plasma-Membrane Calcium Pump: Structure and Function 525
M Brini

Plastocyanin 530
E L Gross

Platelet-Activating Factor Receptor 533
K M Howard

Platelet-Derived Growth Factor Receptor Family 538
M V Kovalenko and A Kazlauskas

Porphyrin Metabolism 544
M Shepherd, A E Medlock, and H A Dailey

Potassium Channels in the Inner Membrane of Mitochondria in Various Organisms: From Unicellular Eukaryotes to Higher Plants and Mammals 550
I Szabò and M Zoratti

Pre-tRNA and Pre-rRNA Processing in Bacteria 554
Z Li

Primary Biliary Cirrhosis 561
M M Kaplan

Prions of Yeast and Fungi: Proteins Acting as Genes 566
R B Wickner

Prions Overview 571
O Bannach and D Riesner

Processivity Clamps in DNA Replication 576
N Y Yao and M O'Donnell

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Prostaglandins and Leukotrienes 581
W L Smith and R C Murphy

Proteases in Blood Clotting 585
J D Kulman and E W Davie

Proteasomes, Overview 590
M Rechsteiner

26S Proteasome: Structure and Function 595
F Förster and E Sakata

Proteinase-Activated Receptors 601
R Ramachandran, K K Hansen, and M D Hollenberg

Protein Carboxyl Esterification 607
E Perez and J B Stock

Protein Data Resources 611
N Furnham, J M Thornton, and J S Garavelli

Protein Degradation 617
A L Goldberg

Protein Folding and Assembly 625
D P Goldenberg

Protein Import into Mitochondria 632
J M Herrmann and W Neupert

Protein Kinase C Family 637
A C Newton

Protein N-Myristoylation 641
E D Sonnenburg and J I Gordon

Protein Palmitoylation 645
R J Deschenes

Protein Tyrosine Phosphatases 648
D J Pagliarini, F L Robinson, C A Worby, and J E Dixon

Proteoglycans 654
S K Olson and J D Esko

Proton Pumping in the Respiratory Chain 661
V Rauhamäki and M Wikström

Pteridines 666
S C Daubner and P F Fitzpatrick

P-Type Pumps: Copper Pump 670
Y Hatori and S Lutsenko

P-Type Pumps: H^+/K^+ Pump 676
J M Shin and G Sachs

P-Type Pumps: Na^+,K^+-ATPase 681
H Haviv and S J D Karlish

P-Type Pumps: Plasma-Membrane H^+ Pumps 688
A B Mason and C W Slayman

Purple Bacteria: Electron Acceptors and Donors 693
A Adessi and R De Philippis

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Purple Bacteria: Photosynthetic Reaction Centers 700
C R D Lancaster

Pyridoxal Phosphate 708
M D Toney

Pyrimidine Biosynthesis and Degradation (Catabolism) 712
M Löffler and E Zameitat

Pyruvate Kinase 719
K Uyeda

Quinones 722
G Lenaz and M L Genova

VOLUME 4

Rab Family 1
M W McCaffrey and A J Lindsay

Ran GTPase 7
M S Moore

Ras Family 12
L A Quilliam

Reactive Oxygen and Nitrogen Species: Interactions with Mitochondria and Pathophysiology 17
B R Zelickson, S W Ballinger, L J Dell'Italia, J Zhang, and V M Darley-Usmar

Recombination-Dependent DNA Replication 23
K N Kreuzer

Recombination: DNA Helicases and Nucleases 29
J Sowa and G Ira

Recombination: DNA-Strand Transferases 35
W D Wright and W-D Heyer

RecQ Helicase Systems 43
G Mosedale and I D Hickson

Regulated Intramembrane Proteolysis 50
J Ye

Regulation of Body Weight by Malonyl-CoA in the CNS 56
S Rodriguez and M J Wolfgang

Regulation of Chromatin Dynamics 59
E-J E Hong and C L Peterson

Regulation of Gene Transcription by Hypoxia-Inducible Factor 1 67
G L Semenza

Renewable Hydrogen from Biomass 72
S Thammannagowda, L Magnusson, J H Jo, P-C Maness, and M Seibert

Repair of G–T Mismatches by *Escherichia coli* Vsr and Eukaryotic DNA Glycosylases 76
M Lieb and A S Bhagwat

Respiration in Phototrophic Microorganisms 79
R Borghese and D Zannoni

Respiratory Chain and ATP Synthase 83
D G Whitehouse and A L Moore

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Respiratory Chain Complex I 87
U Brandt and V Zickermann

Respiratory Chain Complex II and Succinate:Quinone Oxidoreductases 92
C R D Lancaster

Respiratory Chain Complex IV 98
H Michel

Retinoblastoma Protein (pRB) 104
N Singh, M Frolov, and N Dyson

Retinoic Acid Receptors 108
M Petkovich

Rho GTPases and Actin Cytoskeleton Dynamics 111
P Villalonga and A J Ridley

Ribosome Assembly 116
A Sahasranaman and J L Woolford Jr.

Ribosome Regulation by EF-G and EF-Tu 122
S T Gregory

Ribosome Structure 128
T M Schmeing

Riboswitches 136
A R Ferré-D'Amaré

Ribozymes and Evolution 142
N Lehman

Ribozyme Structural Elements: Group II Introns and the Spliceosome 147
O Fedorova and C Waldsich

RNA Editing 154
C E Samuel

RNA Polymerase I and RNA Polymerase III in Eukaryotes 159
R J White

RNA Polymerase II and Its General Transcription Factors 162
J L Corden

RNA Polymerase II Elongation Control in Eukaryotes 164
D H Price

RNA Polymerase Reaction in Bacteria 167
S-H Jun, B A Warner, and K S Murakami

RNA Polymerase Structure, Bacterial 173
S Borukhov

RNA Splicing 185
M S Jurica and G A Roybal

Roles of Micro-RNAs in Metabolism 191
Z Wei and G W Wong

Ryanodine Receptor Calcium Ion Channels 195
G Meissner

Secretases 199
R L Heinrikson

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Secretory Pathway 203
S Shikano and K J Colley

Selenoprotein Synthesis 210
A Böck

Septins and Cytokinesis 214
A S Maddox and C M Field

Serine/Threonine Phosphatases 219
T S Ingebritsen

Serotonin Receptor Signaling 224
P J Gresch

Siglecs 229
A Varki

Sigma Factors 232
J D Helmann and S R MacLellan

Sliding Clamps in DNA Replication: *Escherichia coli* β-Clamp and PCNA Structure 236
M R Marzahn and L B Bloom

Small GTPases 242
K H Pedone and C J Der

Small RNAs in Bacteria 249
E R Westra, M M Jore, S Al-Attar, S J J Brouns, and J van der Oost

Sodium–Calcium Exchanger: Structural Aspects 256
M Ottolia, D A Nicoll, and K D Philipson

Somatostatin Receptors 262
A Schonbrunn

Spastic Paraplegia 267
E I Rugarli

Spectrophotometric Assays 273
B Chance

Sphingolipid Biosynthesis 281
A H Merrill Jr.

Sphingolipid Catabolism 287
A Abe and J A Shayman

Src Family of Protein Tyrosine Kinases 293
J A Cooper

Starvation 298
O E Owen and R W Hanson

Steroid/Thyroid Hormone Receptors 308
R D Ward and N L Weigel

Store-Operated Calcium Channels 314
A Amcheslavsky, A V Yeromin, A Penna, and M D Cahalan

Structure and Regulation of Pyruvate Dehydrogenases 321
J L S Milne

Structure–Function of the Cytochrome b_6f Complex of Oxygenic Photosynthesis 329
W A Cramer, E Yamashita, D Baniulis, J Whitelegge, and S S Hasan

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Structure of P-Type Adenosine Triphosphatases 335
I Dach and P Nissen

Structure of Respiratory Complex I 341
J M Berrisford and L A Sazanov

Substrate Binding, Catalysis, and Product Release 347
W W Cleland

Sugar Nucleotide Transporters 350
C B Hirschberg

Superoxide Dismutase 352
I Fridovich

T7 RNA Polymerase 355
R Sousa

Tachykinin/Substance P/Neurokinin-1 Receptors 360
L E Goldsmith and M M Kwatra

Taste Receptors 366
J D Boughter Jr. and S D Munger

T-Cell Antigen Receptor 369
A L Szymczak-Workman, L L Jones, C S Guy, and D A A Vignali

T-Cell Receptor Signaling to NF-κB 376
R R McCully and J L Pomerantz

Telomeres: Maintenance and Replication 380
A Bianchi and D Shore

Thyroid-Stimulating Hormone/Luteinizing Hormone/Follicle-Stimulating Hormone Receptors 387
D L Segaloff and M Ascoli

Tight Junctions 392
S Tsukita

Toll-Like Receptors 396
H Kumar, O Takeuchi, and S Akira

Transcription-Coupled DNA Repair Overview 402
S Tornaletti

Transcription Termination 407
T J Santangelo

Transient Receptor Potential Channels 412
I S Ambudkar, X Liu, and B Paria

Translation Elongation in Bacteria 418
S C Blanchard

Translation Initiation in Bacteria: Factors and Mechanisms 425
L Brandi and C O Gualerzi

Translation Initiation in Eukaryotes: Factors and Mechanisms 432
T V Pestova and C U T Hellen

Tricarboxylic Acid Cycle 436
R L Veech

tRNA Synthetases 441
R W Alexander

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Troponin 445
I Ohtsuki and S Morimoto

Tubulin and Its Isoforms 450
E Nogales

Tumor Necrosis Factor Receptors 454
T C Cheung and C F Ware

Two-Hybrid Protein–Protein Interactions 460
I Serebriiskii, S Peri, and E A Golemis

Tyrosine Sulfation 463
D Corbeil and W B Huttner

Ubiquitin-Like Proteins 467
L-S Lu and E T H Yeh

Ubiquitin System 473
A Ciechanover and M H Glickman

UmuC D Lesion Bypass DNA Polymerase V 477
L A Hawver and P J Beuning

Uncoupling Proteins 482
D Ricquier and F Bouillaud

Unfolded Protein Responses 488
D Ron

Urea Cycle: Disease Aspects 494
C Ficicioglu and M Yudkoff

Vacuoles 502
C J Stefan, W M Henne, and S D Emr

Vascular Endothelial Growth Factor Receptors 509
K A Thomas

Vasopressin/Oxytocin Receptor Family 515
M J Brownstein

V-ATPases 520
J Capecci and M Forgac

Vitamin A (Retinoids) 525
J L Napoli

Vitamin C 530
F M Steinberg and R B Rucker

Vitamin D 535
H F DeLuca and M Clagett-Dame

Vitamin D Receptor 540
D R Dowd and P N MacDonald

Vitamin E 545
J Atkinson, D Manor, and R Parker

Vitamin K: Biochemistry, Metabolism, and Nutritional Aspects 551
J W Suttie

Voltage-Dependent K^+ Channels 555
R Latorre, F J Morera, and C Zaelzer

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Voltage-Gated Ca^{2+} Channels 560
H Reuter

Voltage-Gated Sodium Channels: Structure, Function, and Pathophysiology 564
W A Catterall

von Hippel–Lindau (VHL) Protein 570
R C Conaway and J W Conaway

XPV Polymerase and the Bypass of Ultraviolet DNA Damage 572
A R Lehmann

Zinc Fingers 575
M Isalan

Index *581*

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

This material may be protected by Copyright law (Title 17 U.S. Code)

Heme Synthesis

G C Ferreira, University of South Florida, Tampa, FL, USA

© 2013 Elsevier Inc. All rights reserved.

Glossary

Erythroblast Erythroid cell precursor; nucleated cell in bone marrow that develops into an erythrocyte.
Erythrocyte A red blood cell; non-nucleated, disk-shaped blood cell that contains hemoglobin.
Erythroid Erythrocyte related.
Erythropoiesis Formation and production of erythrocytes.
Tetrapyrrole A general term that refers to molecules with four rings of the pyrrole type, most often linked together by single-atom bridges between the α-positions of the five-membered pyrrole rings. The arrangement of the four rings is macrocyclic in porphyrins.

Introduction

Heme, a four-ring organic compound with a central iron atom, plays multiple roles in cellular processes. For example, it is the crucial component in hemoglobin, a red blood cell protein responsible, in vertebrates, for the transport of oxygen from the lungs to the different tissues and the transport of carbon dioxide from the tissues to the lungs. It is heme that gives the characteristic red color to blood. Given its importance in many vital functions in organisms ranging from bacteria to man, it is not surprising that most cells synthesize heme. However, in humans, differentiating erythrocytes represent the major site of heme production due to the synthesis of hemoglobin, which accounts for ~85% of the total heme.

Heme Structure

The observation that hemoglobin can be split into two components, with the smaller unit corresponding to the compound which today is recognized as heme, was made by L R LeCanu in 1837. While the correct structure of heme was initially proposed by Kuster in 1912, it only received general acceptance 16 years later upon its chemical synthesis. The Nobel prize-winning organic chemist Hans Fischer developed syntheses leading to the assembly of four rings, or pyrroles, into a macrocyclic structure with a purple–red color, which he appropriately named 'porphyrin' (from the Greek porphuros, meaning purple). The chemically synthesized porphyrin, which had properties identical to those of the natural porphyrin (derived from heme upon the release of the iron atom), was designated protoporphyrin IX. Thus, heme consists of four rings (pyrroles), composed of carbon, nitrogen, and hydrogen atoms, linked together in a cyclic array and containing an atom of iron in its center (Figure 1).

Heme is now recognized as a member of a large family of metalated tetrapyrroles, which are essential for life on Earth. They include chlorophylls, cobalamin (vitamin B_{12}), siroheme, and coenzyme F_{430}. Chlorophylls trap the light energy required in photosynthesis. Cobalamin-containing enzymes catalyze intramolecular rearrangements, methylations, and reduction of ribonucleotides to deoxyribonucleotides. Siroheme is an essential component of enzymes involved in sulfur and nitrogen metabolisms. Coenzyme F_{430}, present in methanogenic bacteria, functions in methane formation. While the central metal ions in these tetrapyrroles differ – Fe^{2+} in the case of heme and siroheme, Mg^{2+} in chlorophyll and bacteriochlorophyll, Co^{2+} in cobalamin, and Ni^{2+} in coenzyme F_{430} – the unifying feature is the presence of a complexed metal ion, which promotes a remarkable number of different biochemical reactions. Further, the protein scaffold binding the metalloporphyrin modulates the chemistry (redox state and coordination properties) of the metal ion, such that the spectrum of the reactions expands beyond the normal properties of the metal ion. It is this interplay between the protein moiety and the metalloporphyrin that provides the plethora of functions associated with metalloporphyrin-containing proteins.

The Heme Biosynthetic Pathway

Formation of 5-Aminolevulinate

The biosynthesis of tetrapyrroles is initiated from simple molecules and comprises only a few enzymatic steps (Figure 2). The first committed precursor, 5-aminolevulinic acid (ALA), is synthesized from either succinyl-CoA or glutamate, and thus the names C4 and C5 (based on the number of carbon atoms in the precursor), for the two ALA biosynthetic pathways evolved in nature. In most nonphotosynthetic eukaryotes and a few prokaryotes (i.e., the α-proteobacteria), ALA results from the condensation between glycine and succinyl CoA (C4 pathway), while in plants and most bacteria, ALA is derived from glutamate (C5 pathway). The C4 pathway requires a single enzyme, ALA synthase (ALAS); by contrast, three enzymes are involved in the conversion of glutamate into ALA. Specifically, glutamyl-transfer RNA (tRNA) synthetase converts glutamate to glutamyl-tRNA, which is reduced by glutamyl-tRNA reductase to glutamyl-1-semialdehyde (GSA); this product is then transformed by GSA aminomutase or GSA aminotransferase (GSA-AT) to ALA. Pyridoxal 5′-phosphate, a vitamin B_6-derivative, is an essential component (cofactor) of both ALAS and GSA-AT, as these enzymes cannot function relying solely on the protein moiety. Among prokaryotes, ALAS is only found in the α-proteobacteria, which are considered to be the most closely related, free-living organisms to the

Motif Exhibit 1059, 457 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

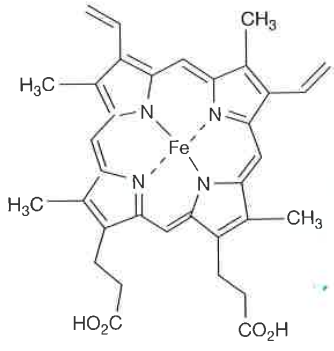


Figure 1 Iron–protoporphyrin IX, which is commonly designated as heme. Iron–protoporphyrin IX (or heme) has an iron ion complexed with the conjugated, aromatic porphyrin ring and an asymmetrical side-chain substitution. Note the $-CH_3$ (methyl), $-CH=CH_2$ (vinyl), and $-CH_2-CH_2-COOH$ (propionate) groups as side chains.

ancestral progenitor of the mitochondrion. Thus, it is conceivable that eukaryotic ALASs were acquired from a free-living α-proteobacterium during its transformation, first to an endosymbiont and subsequently to a mitochondrion.

From ALA to Uroporphyrinogen III

In eukaryotes, the first and three final enzymatic steps of heme biosynthesis occur within mitochondria, while the intervening four occur in the cytosol (Figure 2). Once ALA is made and transported to the cytosol, through most likely a specialized mitochondrial transporter, two molecules of ALA are asymmetrically condensed to form the monopyrrole porphobilinogen (PBG) by PBG synthase (PBGS; also known as ALA dehydratase). PBG constitutes the building unit of tetrapyrroles and hence eight molecules of ALA are required to form one heme molecule. Crystal structures of PBGSs from different species have confirmed the presence of two distinct binding sites for each ALA molecule, termed A and P to distinguish the two ALA substrate molecules. The A-side of ALA forms the acetyl half of PBG, leaving its amino group intact, while the P-side of ALA

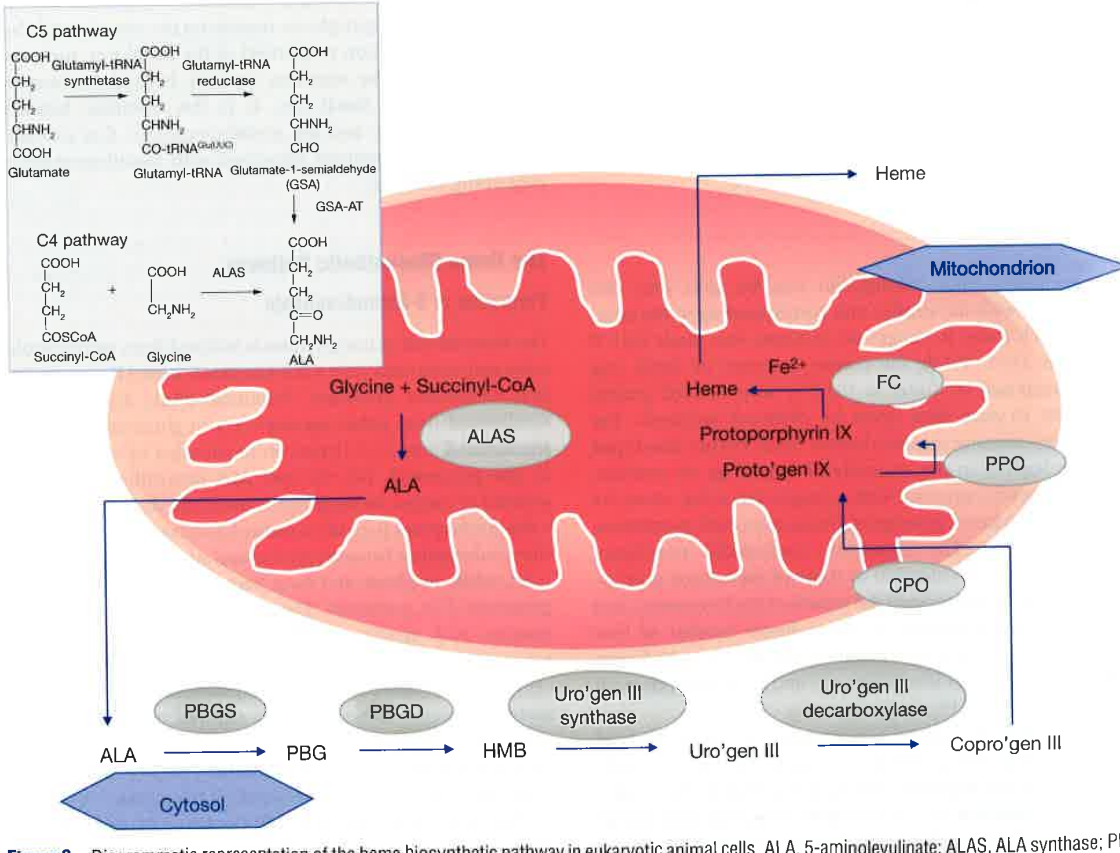


Figure 2 Diagrammatic representation of the heme biosynthetic pathway in eukaryotic animal cells. ALA, 5-aminolevulinate; ALAS, ALA synthase; PBG, porphobilinogen; PBGS, PBG synthase; PBGD, PBG deaminase; HMB, hydroxymethylbilane; Uro'gen III, uroporphyrinogen III; copro'gen III, coproporphyrinogen III; CPO, coproporphyrinogen III oxidase; proto'gen IX, protoporphyrinogen IX; PPO, protoporphyrinogen IX oxidase; FC, ferrochelatase. Inset: The two paths to ALA.

Motif Exhibit 1059, 458 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

forms the propionyl half of PBG with the amino group becoming part of the PBG ring. PBGS has a homo-octomeric structure composed of four dimers, and each subunit contains an active site metal ion (zinc or magnesium and/or potassium depending on the species) and in many cases an allosteric metal ion (magnesium). The next step in heme biosynthesis involves the deamination and stepwise polymerization of four PBG molecules by PBG deaminase to give the linear tetrapyrrole 1-hydroxymethylbilane. Then, uroporphyrinogen III synthase catalyzes the cyclization of the linear tetrapyrrole, whereby the D ring of hydroxymethylbilane is inverted to generate the asymmetrical macrocycle, uroporphyrinogen III. In the absence of uroporphyrinogen III synthase, 1-hydroxymethylbilane cyclizes into the nonphysiological relevant tetrapyrrole, uroporphyrinogen I. The incredible acrobatics that must occur with the flipping of the D ring and closure of hydroxymethylbilane speak for the importance of uroporphyrinogen III synthase in making uroporphyrinogen III synthesis reaction. In fact, disastrous consequences emerge with the malfunction of uroporphyrinogen III synthase, as observed in patients with the blood disorder congenital erythropoietic porphyria, who typically have high levels of uroporphyrinogen I and derivatives thereof.

From Uroporphyrinogen III to Protoporphyrinogen IX

For most eukaryotes, the last cytosolic step is the sequential decarboxylation of the four acetate side chains of uroporphyrinogen III to methyl groups by uroporphyrinogen decarboxylase (Figure 2). This enzyme is a $(\alpha/\beta)_8$ barrel protein, which appears to be functional as a homodimer and not to require prosthetic groups or cofactors. Remarkably, uroporphyrinogen decarboxylase enhances the rate of substrate decarboxylation by a factor of 1.2×10^{17}, one of the largest rate enhancements reported for an enzyme that functions without the help of cofactors. The uroporphyrinogen decarboxylase-catalyzed reaction gives coproporphyrinogen III, which is the substrate of coproporphyrinogen oxidase (CPO). The oxidative decarboxylation of the propionate side chains of the rings A and B of coproporphyrinogen III yields protoporphyrinogen IX. While eukaryotic CPO activity depends on oxygen, prokaryotic organisms, under anaerobic conditions, use alternative electron acceptors in the oxidative decarboxylation reaction. Hence, oxygen-independent CPOs (HemN and HemZ, i.e., the products of the *hemN* and *hemZ* genes) are found in bacteria; indeed, facultatively anaerobic bacteria, such as *Escherichia coli*, possess both oxygen-dependent CPO (HemF) and oxygen-independent CPO (HemN). HemN contains an iron–sulfur center ([4Fe–4S]), which is essential for the activity of the enzyme in conjunction with nicotinamide adenine dinucleotide (phosphate), reduced (NAD(P)H) and *S*-adenosyl-L-methionine (SAM). As one would expect, the enzyme appears to have a distinct reaction mechanism with the involvement of SAM in the formation of a radical, and thus HemN has been aptly identified as a radical SAM enzyme. As far as the subcellular location of the oxygen-dependent CPO goes, it seems to vary! In mammals and other higher eukaryotes, it appears to be associated with the inner side of the mitochondrial outer membrane, whereas in the yeast *Saccharomyces cerevisiae* and bacteria it is found in the cytosol.

From Protoporphyrinogen IX to Heme

In the penultimate step, protoporphyrinogen oxidase (PPO) catalyzes the complete aromatization of protoporphyrinogen IX ring into protoporphyrin IX, which involves a six-electron oxidation (Figure 2). Similar to CPO, anaerobic and facultative anaerobic bacteria have an oxygen-independent PPO; however, in contrast to CPO, facultative anaerobic bacteria possess solely an oxygen-independent PPO (and not in complementation with an oxygen-dependent PPO). All other organisms appear to have PPOs, which use oxygen as the terminal electron acceptor. Ferrochelatase catalyzes the terminal step of the heme biosynthetic pathway, the insertion of ferrous iron into the protoporphyrin IX macrocyclic ring. Eukaryotic ferrochelatases associate with the inner side of the mitochondrial inner membrane, and vertebrate enzymes have a [2Fe–2S] cluster (although this is not an exclusive structural feature of vertebrate ferrochelatases). The mechanism of the ferrochelatase reaction involves an out-of-plane distortion of the porphyrin, which facilitates the reaction by exposing the pyrrole nitrogens to the incoming ferrous iron substrate.

Regulation of the Heme Biosynthetic Pathway/Heme Biosynthesis

There is not a single ubiquitous regulatory mechanism for heme biosynthesis, and so the discussion in this article focuses solely on the regulation of the pathway in mammalian cells. Considering that heme is predominantly synthesized in bone marrow (i.e., in erythroblasts and reticulocytes that still contain mitochondria) and liver cells, most of the studies on regulation have focused on these two types of cells. Predominance of heme synthesis in bone marrow and liver cells reflects the greater demands due to the syntheses of the heme-containing proteins, hemoglobin, and cytochrome P-450 enzymes, respectively. The major regulatory and rate-limiting step of the pathway corresponds to the reaction catalyzed by ALAS, the first enzyme of the pathway. Two genes, located in two different chromosomes in humans, encode the two ALAS forms of the enzyme (or isozymes): ALAS1 and ALAS2. The gene encoding ALAS1 is expressed in every cell, whereas that for ALAS2 is specifically expressed in erythroid cells. Importantly, the regulatory mechanisms for the expression of the two ALAS genes and the posttranscriptional events seem to be distinct in differentiating erythrocytes from those in nonerythroid cells. ALAS1 controls the production of heme for basic cellular functions and is feedback regulated by heme through (1) repression of transcription of the ALAS1 gene, (2) inhibition of translation of the ALAS1 message, (3) destabilization of the ALAS1 message, and (4) inhibition of import of the cytosolic precursor form of the enzyme into mitochondria. By contrast, certain drugs and hormones induce liver cells to make more ALAS, and consequently heme and cytochrome P-450, by activating transcription of the ALAS1 gene.

In differentiating erythroid cells, the syntheses of heme and globin need to be coordinated, as the production of hemoglobin requires that neither the protein (globin) component nor the nonprotein component (heme) will be in excess. And indeed, identical transcriptional elements have been identified

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

in the genes of ALAS2 and globin. Another level of regulation of the erythroid heme biosynthetic pathway is centered at the stage of translation, which responds to the cellular iron levels. This response is mediated through specific interactions between an iron-responsive element (IRE), present in the ALAS2 messenger RNA (mRNA), and IRE-binding proteins, such that under iron limitation, the translation of ALAS2 mRNA is inhibited, while the presence of iron above the cellular threshold relieves the repression caused by IRE-binding proteins. The essential requirement for ALAS2 in heme biosynthesis and erythropoiesis was clearly demonstrated with targeted disruption of ALAS2 in the mouse, which led to no hemoglobin in cells, the accumulation of iron (albeit in the cytoplasm and not in mitochondria) and, ultimately, embryonic death.

Heme Function

The known roles of heme are many and variable, and most probably many other functions await discovery. As referred to above, heme is a prosthetic group of proteins involved in oxygen transport such as hemoglobins and myoglobins, which are the major blood and muscle proteins, respectively. In addition, heme participates in other cellular processes including anaerobic and aerobic respiration, sensing of diatomic gases (e.g., nitric oxide and carbon monoxide), detoxification of foreign substances (e.g., drugs) by oxidative metabolism, and signal transduction. Some of these functions might also be coupled with the regulatory roles of heme in gene transcription, translation, microRNA (miRNA) processing, mitochondrial protein import, protein stability, and differentiation. A transcription factor that regulates the mammalian biological clock was also discovered to use heme as a cofactor to sense, and react to, changes in the cell, thus linking cellular metabolism with the biological clock. Another area of research has led to the emergence of a new proposal that heme deficiency or dysregulation is intimately associated with the aging process, including neural decay.

Disorders of the Heme Biosynthetic Pathway

Malfunction of the heme biosynthetic pathway can have disastrous consequences, and in humans it can be manifested in porphyrias. Briefly, porphyrias are disorders that result from inherited or acquired defects in the heme biosynthetic pathway. The first reports of porphyrias were in 1874. Since then, eight types of porphyrias associated with defects in each of the eight enzymes of the pathway have been identified (Table 1), and depending on the major site of expression of the enzymatic defect, they have been classified into hepatic or erythropoietic porphyria. Although the molecular mechanisms behind the clinical manifestations are not well understood, the two major symptoms include cutaneous photosensitivity and central nervous system disturbances, and thus the other classification of either photosensitive or neurological porphyrias. Some porphyrias, however, manifest overlapping symptoms. Photosensitivity is a result of porphyrin accumulation, which typically occurs when the defective enzyme is anywhere in the pathway beyond the third step of the pathway (i.e., the PBG deaminase-catalyzed reaction), whereas the neurologic manifestations are associated with overproduction of ALA and PBG. However, the first porphyria associated with erythroid ALAS, named X-linked dominant protoporphyria (XLDPP), was recently identified, and elevated concentration of ALA-derived protoporphyrin IX and Zn-protoporphyrin were observed in the erythrocytes of the XLDPP patients. Since the 1980s, many of the genetic mutations responsible for the enzymatic defects have been identified and mouse models, targeted to correct the defective enzyme, have been developed by engineering functional enzyme expression in bone marrow cells. Zebrafish have also provided powerful molecular models for erythropoietic porphyrias. For example, mutations in the ferrochelatase gene resulted in photosensitive zebrafish with autofluorescent blood. Finally, mutations in the gene encoding for the erythroid ALAS can cause X-linked sideroblastic anemia (Table 1), a blood disorder characterized by mitochondrial iron accumulation due to a reduction in heme biosynthesis destined for hemoglobin production in erythroblasts.

Table 1 Heme biosynthesis-associated disorders

Disorder	*Defective enzyme*
X-linked sideroblastic anemia	5-Aminolevulinate synthase (erythroid)
X-linked dominant protoporphyria	5-Aminolevulinate synthase (erythroid)
ALA dehydratase, or Doss, porphyria	Porphobilinogen synthase (or ALA dehydratase)
Acute intermittent porphyria	Porphobilinogen deaminase
Congenital erythropoietic porphyria	Uroporphyrinogen III synthase
Porphyria cutanea tarda	Uroporphyrinogen decarboxylase
Coproporphyria	Coproporphyrinogen oxidase
Variegate porphyria	Protoporphyrinogen oxidase
Erythropoietic protoporphyria	Ferrochelatase

See also: **Metabolism Vitamins and Hormones:** Porphyrin Metabolism; **Protein/Enzyme Structure Function and Degradation:** Heme Proteins; **Signaling:** Hematopoietin Receptors.

Further Reading

Ajioka RS, Phillips JD, and Kushner JP (2006) Biosynthesis of heme in mammals. *Biochimica et Biophysica Acta* 1763: 723–736.

Bottomley SS (2008) Sideroblastic anaemias. In: Greer JP, Foerster J, Rodgers GM, et al. (eds.) *Wintrobe's Clinical Hematology*, 12th edn., pp. 835–873. Philadelphia, PA: Lippincott Williams and Wilkins.

Ferreira GC (1999) 5-Aminolevulinate synthase and mammalian heme biosynthesis. In: Ferreira GC, Moura JG, and Franco R (eds.) *Iron Metabolism. Inorganic Biochemistry and Regulatory Mechanism*, pp. 15–34. Weinheim, Germany: Wiley-VCH.

Ferreira GC and Hunter GA (2011) Ferrochelatase structure and reaction mechanism. In: Kadish KM, Smith KM, and Guilard R (eds.) *Handbook of Porphyrin Science*, pp. 49–121. Hackensack, NJ: World Scientific Publishing.

Latunde-Dada GO, Simpson RJ, and McKie AT (2006) Recent advances in mammalian haem transport. *Trends in Biochemical Sciences* 31: 182–188.

Severance S and Hamza I (2009) Trafficking of heme and porphyrins in metazoa. *Chemical Reviews* 109: 4596–4616.

Motif Exhibit 1059, 460 of 653

Ex. 1041 Cantor 1999

GENOMICS

The Science and Technology Behind the Human Genome Project

Charles R. Cantor
Cassandra L. Smith
Center for Advanced Biotechnology
Boston University
Boston, Massachusetts



A Wiley-Interscience Publication

JOHN WILEY & SONS, INC.

New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

Steer[illegible]
Univers[illegible] Madison
550 [illegible] Drive
Madison, WI 53706-1293

This book is printed on acid-free paper. ♾

Copyright © 1999 by John Wiley & Sons, Inc. All rights reserved.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4744. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, (212) 850-6011, fax (212) 850-6008, E-Mail: PERMREQ@WILEY.COM.

Library of Congress Cataloging-in-Publication Data
Cantor, Charles R.
Genomics : the science and technology behind the human genome project / Charles R. Cantor and Cassandra Smith.
p. cm.
Includes index.
ISBN 0-471-59908-5 (alk. paper)
1. DNA--Analysis. 2. Nucleotide sequence. 3. Gene mapping.
I. Smith, Cassandra. II. Title.
QP624.C36 1999
572.8' 6--dc21 98-40448

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

CONTENTS

Preface xiii

Introduction xv

1 DNA Chemistry and Biology 1

Basic Properties of DNA 1
Covalent Structure 1
Double Helical Structure 1
Methylated Bases 5
Plasticity in DNA Structure 7
DNA Synthesis 8
DNA as a Flexible Set of Chemical Reagents 15
Basic DNA Biology 19
Genome Sizes 25
Number of Genes 27
Sources and Additional Readings 27

2 A Genome Overview at the Level of Chromosomes 29

Basic Properties of Chromosomes 29
Bacterial Chromosomes 29
Chromosomes of Eukaryotic Organisms 32
Centromeres 32
Telomeres 34
Dynamic Behavior of Telomeres 35
Chromatin and the Higher-Order Structure of Chromosomes 36
Chromosomes in the Cell Cycle 39
Genome Organization 40
Chromosome Purification 43
Chromosome Number 51
Unusual Characteristics of Sex Chromosomes and Mitochondria 54
Synteny 59
Sources and Additional Readings 63

3 Analysis of DNA Sequences by Hybridization 64

Basic Requirements for Selectivity and Sensitivity 64
Detection of Specific DNA Sequences 66
Equilibria between DNA Double and Single Strands 67
Thermodynamics of the Melting of Short Duplexes 71

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Thermodynamics of Imperfectly Paired Duplexes 74
Kinetics of the Melting of Short Duplexes 77
Kinetics of Melting of Long DNA 79
Kinetics of Double-Strand Formation 80
Complexity 85
Hybridization on Filters 86
Sensitive Detection 90
Sources and Additional Readings 97

4 Polymerase Chain Reaction and Other Methods for In Vitro DNA Amplification **98**

Why Amplify DNA? 98
Basic Principles of the Polymerase Chain Reaction (PCR) 98
Noise in PCR: Contamination 103
PCR Noise: Mispriming 104
Misincorporation 106
Long PCR 106
Incorporating Extra Functionalities 107
Single-Sided PCR 107
Reducing Complexity with PCR 112
Additional Variants of the Basic PCR Reaction 114
Total Genome Amplification Methods 116
Application of PCR to Detect Molecules Other Than DNA 119
DNA Amplification without Thermal Cycling and Other Alternatives to PCR 122
Future of PCR 127
Sources and Additional Readings 128

5 Principles of DNA Electrophoresis **131**

Physical Fractionation of DNA 131
Separation of DNA in the Ultracentrifuge 131
Electrophoretic Size Separations of DNA 132
Electrophoresis without Gels 133
Motions of DNA Molecules in Gels 135
Complex Effects of Gel Structure and Behavior 136
Biased Reptation Model of DNA Behavior in Gels 138
Pulsed Field Gel Electrophoresis (PFG) 140
Macroscopic Behavior of DNA in PFG 146
Inadequacy of Reptation Models for PFG 148
DNA Trapping Electrophoresis 155
Secondary Pulsed Field Gel Electrophoresis (SPFG) 157
Entry of DNAs into Gels 158
Sources and Additional Readings 164

6 Genetic Analysis **165**

Why We Need Genetics 165
Basic Strategy for Genetic Analysis in the Human: Linkage Mapping 165

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

A Glossary of Genetic Terms 170
Relationship between the Physical and the Genetic Maps 174
Power of Mouse Genetics 178
Weakness of Human Genetics 178
Linkage Analysis Ignoring Recombination 180
Linkage Analysis with Recombination 183
Interval Mapping 185
Finding Genes by Genetic Mapping 188
Moving from Weak Linkage Closer to a Gene 190
Linkage Disequilibrium 191
Complications in Linkage Disequilibrium and Genetic Maps in General 193
Distortions in the Genetic Map 194
Current State of the Human Genetic Map 195
Genetics in the Pseudoautosomal Region 197
Why Genetics Needs DNA Analysis 201
Detection of Homozygous Regions 204
Sources and Additional Readings 206

7 Cytogenetics and Pseudogenetics 208

Why Genetics Is Insufficient 208
Somatic Cell Genetics 208
Subchromosomal Mapping Panels 210
Radiation Hybrids 212
Single-Sperm PCR 215
In Situ Hybridization 218
High-Resolution FISH 224
Chromosome Painting 229
Chromosome Microdissection 230
Sources and Additional Readings 232

8 Physical Mapping 234

Why High Resolution Physical Maps Are Needed 234
Restriction Maps 235
Ordered Libraries 239
Restriction Nuclease Genomic Digests 241
HTF Islands 245
Ordering Restriction Fragments 246
Identifying the DNA Fragments Generated by a Rare-Cutting Restriction Enzyme 248
Mapping in Cases Where Fragment Lengths Can Be Measured Directly 252
Generation of Larger DNA Fragment Sizes 253
Linking Clones 254
Jumping Libraries 257
Partial Digestion 259
Exploiting DNA Polymorphisms to Assist Mapping 262
Placing Small Fragments on Maps 264

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Reaching the Ends of the Physical Map: Cloning Telomeres 265
Optical Mapping 269
Bottom-Up Library Ordering 269
Measurements of Progress in Building Ordered Libraries 275
Survey of Restriction Map and Ordered Library Construction 277
Sources and Additional Readings 284

9 Enhanced Methods for Physical Mapping **285**

Why Better Mapping Methods Are Needed 285
Larger Yeast Artificial Chromosomes (YACs) 285
How Far Can YACs Go? 288
Vector Obsolescence 290
Hybrid Mapping Strategies: Cross-connections between Libraries 291
Screening by PCR versus Hybridization 296
Tiered Sets of Samples 298
Simple Pooling Strategies for Finding a Clone of Interest 300
Sequence-Specific Tags 301
Pooling in Mapping Strategies 303
Probe Pooling in *S. pombe* Mapping 305
False Positives with Simple Pooling Schemes 311
More General Pooling Schemes 312
Alternate Array Configurations 316
Inner Product Mapping 318
Sliced PFG Fractionations as Natural Pools of Samples 320
Restriction Landmark Genome Scanning 320
Prognosis for the Future of Genome Mapping 322
Sources and Additional Readings 323

10 DNA Sequencing: Current Tactics **325**

Why Determine DNA Sequence 325
Design of DNA Sequencing Projects 326
Ladder Sequencing Tactics 327
Issues in Ladder Sequencing 330
Current Fluorescent DNA Sequencing 334
Variations in Contemporary DNA Sequencing Tactics 336
Errors in DNA Sequencing 341
Automated DNA Sequencing Chemistry 345
Future Improvements in Ladder Sequencing 348
Approaches to DNA Sequencing by Mass Spectrometry 349
Rate-Limiting Steps in Current DNA Sequencing 358
Sources and Additional Readings 359

11 Strategies for Large-Scale DNA Sequencing **361**

Why Strategies Are Needed 361
Shotgun DNA Sequencing 361

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Directed Sequencing with Walking Primers 363
Priming with Mixtures of Short Oligonucleotides 365
Ordered Subdivision of DNA Targets 368
Transposon-Mediated DNA Sequencing 368
Delta Restriction Cloning 370
Nested Deletions 371
Primer Jumping 373
Primer Multiplexing 375
Multiplex Genomic Walking 376
Global Strategies 377
Sequence-Ready Libraries 379
Sequencing cDNA Libraries 380
Dealing with Uneven cDNA Distribution 381
Large-Scale cDNA Sequencing 384
What Is Meant by a Complete Genome Sequence? 389
Sequencing the Fifth Base 390
Sources and Additional Readings 392

12 Future DNA Sequencing without Length Fractionation 394

Why Try to Avoid Length Fractionations? 394
Single-Molecule Sequencing 394
Sequencing by High-Resolution Microscopy 397
Stepwise Enzymatic Sequencing 400
DNA Sequencing by Hybridization (SBH) 403
Branch Point Ambiguities 404
SBH Using Oligonucleotide Chips 406
Sequencing by Hybridization to Sample Chips 410
Early Experiences with SBH 412
Data Acquisition and Analysis 415
Obstacles to Successful SBH 417
SBH in Comparative DNA Sequencing 420
Oligonucleotide Stacking Hybridization 421
Other Approaches for Enhancing SBH 424
Positional Sequencing by Hybridization (PSBH) 425
Combination of SBH with Other Sequencing Methods 430
Sources and Additional Readings 431

13 Finding Genes and Mutations 433

Detection of Altered DNA Sequences 433
Finding Genes 434
Diagnostics at the DNA Level 448
Analysis of DNA Sequence Differences 455
Heteroduplex Detection 456
Diagnosis of Infectious Disease 462
Detection of New Mutations 463
Sources and Additional Readings 467

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

14 Sequence-Specific Manipulation of DNA 470

Exploiting the Specificity of Base-Base Recognition 470
Structure of Triple-Stranded DNA 470
Triplex-Mediated DNA Cleavage 476
Sequence-Specific DNA Capture 480
Triplex-Mediated DNA Capture 480
Affinity Capture Electrophoresis 486
Use of Backbone Analogues in Sequence-Specific DNA Manipulation 489
Sequence-Specific Cloning Procedures 492
Identification or Cloning of Sequences Based on Differences in Expression Level 499
Coincidence Cloning 500
Human Interspersed Repeated DNA Sequences 506
Distribution of Repeats Along Chromosomes 509
PCR Based on Repeating Sequences 510
Repeat Expansion Detection 516
Aptamer Selection Strategies 517
Oligonucleotides as Drugs 520
Sources and Additional Readings 523

15 Results and Implications of Large-Scale DNA Sequencing 526

Costing the Genome Project 526
Finding Genes 530
More Robust Methods for Finding Genes by DNA Sequence Analysis 532
Neural Net Analysis of DNA Sequences 535
Survey of Past Large-Scale DNA Sequencing Projects 540
Finding Errors in DNA Sequences 545
Searching for the Biological Function of DNA Sequences 547
Searching for the Biological Function of Genes 548
Methods for Comparing Sequences 551
Dynamic Programming 557
Gaining Additional Power in Sequence Comparisons 560
Domains and Motifs 561
Interpreting Noncoding Sequence 563
Diversity of DNA Sequences 564
Sources and Additional Readings 565

Appendix: Databases 569
Index 575

PREFACE

This book is an outgrowth of the George Fisher Baker Lecture series presented by one of us (C.R.C.) at Cornell University in the fall of 1992. This author is tremendously grateful to all those at Cornell who made this occasion truly memorable, personally, and most productive, intellectually. Included especially are Jean Fréchet, Barbara Baird, John Clardy, Jerrold Meinwold, Fred McLafferty, Benjamin Widom, David Usher, and Quentin Gibson, among many others. He also wants to express great thanks to Rick Bourke, who, several times provided extraordinary help, without which lectures would have been missed or required rescheduling.

The book also builds in part on the research experience of both authors over the past 15 years. Without the help of numerous talented co-workers and collaborators, we would never have achieved the accomplishments that form some of the work described herein. No book is produced without heroic efforts on the part of support staff who handle endless chores, even in this era of efficient word processing. We are especially grateful to Michael Hirota for hours of patient deciphering of handwritten scribble. A major part of the effort in completing any substantial publication is receiving and dealing with constructive criticism from those who have read early drafts. Joel Graber and Chandran Sabanayagam helped with the illustrations and many other details; Scott Downing, Foual Siddiqi, and many others used preliminary versions as a textbook in courses over the past few years. To them we owe our thanks for taking the time to reveal ambiguities and painful lapses in our logic. Correcting these has, we hope, made the book far more accurate and comprehensible.

Inevitably in a fast-moving field like genome research, parts of this book were out of date even at the time it was submitted for publication. Furthermore, like many books based on a lecture series, this one relies too heavily on the authors' own work to provide a truly fair representation of many parallel or competing efforts in the field. No attempt was made to be comprehensive or representative in citing sources and additional readings. Despite these limitations, by concentration on fundamentals, the authors hope that a vast majority of workers in the field will find this text a useful companion.

CHARLES R. CANTOR
CASSANDRA L. SMITH

Boston Massachusetts

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

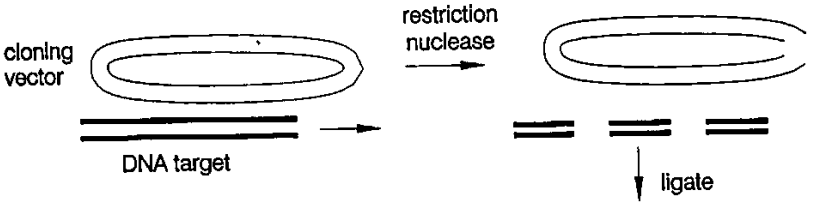
This material may be protected by Copyright law (Title 17 U.S. Code)

BOX 1.3
CLONING DNAs IN BACTERIAL PLASMIDS AND BACTERIOPHAGE

Cloning is the process of making many identical cells (or organisms) from a single precursor. If that precursor contains a target DNA molecule of interest, the cloning process will amplify that single molecule into a whole population. Most single-cell organisms are clonal. Their progeny are identical replicas. Cloning can also be done by manipulating single, immortal, or immortalizable cells of higher organisms including plants and animals. Here we concentrate on bacterial cloning, which can be carried out in a large number of different species but often involves the favorite laboratory organism *Escherichia coli.*

Cloning DNA requires three components: the target DNA of interest, a vector DNA that will accommodate the target as an insertion and that contains a replication origin for it to be propagated indefinitely, and a host cell that will replicate the vector and its insert. The simplest early cloning systems used two types of vectors, either plasmids, which are small DNAs that can replicate independently of the host cell DNA, or bacteriophages, which also have independently replicating DNA systems. Plasmids are usually double-stranded circular DNAs. Bacteriophages can be circular or linear, single stranded, or double stranded. Here we illustrate the steps in cloning double strands.

The vector is usually designed so that it has a unique cutting site for a particular restriction enzyme. Digestion linearizes the vector and produces two ends, usually with the same overhanging sequence. The target is cut with the same enzyme, or alternatively, with a different enzyme that nevertheless leaves the same overhangs. A complex target will yield many fragments, only a few of which may be of interest.

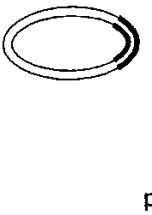


When the vector and target are mixed, they can anneal and form suitable substrates of DNA ligase. Depending on the relative concentrations used, and whether the 5′-phosphate ends produced by the restriction nuclease are left intact or are removed with a phosphatase, a variety of different products combining targets and vectors will be formed. A number of tricks exist to bias the ligation in favor of particular 1:1 target-vector adducts. For example, in the scheme above a single target and vector can come together in two different polarities. More complex schemes allow the polarity to be preselected.

(continued)

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

BOX 1.3 *(Continued)*

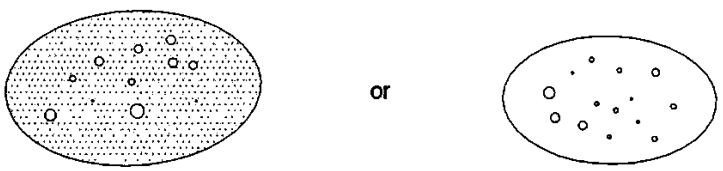


Next the vector-target adduct must be reintroduced into the bacterial cells. Depending on the vector, different procedures are used. If the vector is a plasmid, the bacterial cells are made permeable by chemical, enzymatic, electrical, or mechanical procedures. When mixed with DNA, these bacteria take up some of the vector before the gaps in their surfaces can be resealed. The process is called *transformation* or *transfection*. Very similar procedures can be used to introduce DNA into other kinds of cells. The efficiencies vary widely depending on the type of vector, type of cell, and the procedures used for permeabilization.

If the vector is a bacteriophage, it is usually preferable to repackage the vector-target adduct in vitro and then allow the assembled cells to infect the host cells naturally. This increases the efficiency of delivering DNA into the host cells, both because bacteriophage insertion is often quite effective and because there is no need to permeabilize the host cells, a process that frequently kills many of them.

Traditionally microbiological screening systems are used to detect host cells propagating cloned DNA. If a dilute suspension of bacteria is allowed to coat the surface of a culture dish, colonies will be observed after several days of growth, each of which represents cloned progeny arising from a single cell. The cloning process can be used to introduce markers that aid in the selection or screening of bacteria carrying plasmids of potential interest. For example, the plasmid may carry a gene that confers resistance to an antibiotic so that only host cells containing the plasmid will be able to grow on a culture medium in which that antibiotic is placed. The site at which the vector is cut to introduce the target can disrupt a gene that yields a colored metabolic product when the precursor of that product is present in the medium. Then vectors that have been designated as having no target insert will still yield colored colonies, while those containing a target insert will appear white.

When a bacteriophage vector is used, the same sets of selection and screening systems can be used. If the bacteriophage shows efficient lysis, that is, if infected cells disrupt



(continued)

Motif Exhibit 1059, 472 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

BOX 1.3 *(Continued)*
and release bacteriophage, then a convenient method of cloning is to grow a continuous layer of susceptible bacteria and then cover the culture with a thin layer of dilute bacteriophage suspension. Individual bacteriophage will infect cells, leading to continual cycles of bacteriophage growth and cell lysis until visible holes in the bacterial lawn, called bacteriophage plaques, are observed. Individual colonies or plaques are selected for further study by picking, literally by sampling with a sterile toothpick or other sharp object, and then reinnoculated, respectively, into a permissive growth medium or strain.

A key variable in the choice of vector is the range of target sizes that can be accommodated. Plasmids generally allow a broad range of target sizes, while bacteriophages, because of constraints in the packaging process, generally accept only a narrow range of target sizes. The earliest cloning vectors were efficient for relatively small targets, under 10 kb. Progressively interest has turned to ever larger targets, specialized vector/host systems needed to propagate such targets are described later (Boxes 2.3 and 8.2).

A second key variable in the choice of cloning vector is the copy number. Some plasmids and bacteriophages can grow in tens of thousands of copies per *E. coli* cell. This greatly facilitates recovery and purification of cloned DNA inserts. However, if DNA fragments toxic to the host are the desired targets, a high copy number vector system will make it more difficult for them to be propagated.

The major advantages of cloning, as compared to the in vitro replication systems discussed in Chapter 4, are the ability to handle very large numbers of samples in parallel and the absence of a need for any prior knowledge about the desired target sequences. The major disadvantages of cloning are that it is a relatively time-consuming process difficult to automate, and some DNA sequences are not cloneable either because they are toxic to particular hosts, or they are unstable in those hosts and fragments rearrange or delete.

However, sequences can be designed that associate to give specific structures with three, four, or more ends. These are called DNA junctions (Fig. 1.7). Similar four-stranded structures are actually seen in cells as intermediate steps in genetic recombination. However, these so-called Holliday structures are unstable, since the participating strands are pairs of identical sequences. The location of the junction can move around by a process called *branch migration.* Synthetic junctions can be designed that are stable because their DNA sequences do not allow branch migration. At low salt concentrations these structures become relatively flat, as indicated in Figure 1.7; at higher salt concentrations they form an X-shaped structure with one pair of co-axially stacked helices crossing over a second pair of co-axially stacked helices.

Still other types of DNA structures can be formed by taking advantage of the fact that in the double helix, specific hydrogen bond donor and acceptor sites on the bases remain exposed in the major groove of the double helix. Under appropriate conditions this allows the association of a third DNA strand in a sequence-specific manner to form a DNA triplex. It produces a DNA structure with four ends. The great utility of triplexes, which we will exploit in considerable detail later (Chapter 14), is that triplexes can be formed

Ex. 1042 Alberts 2002

MOLECULAR BIOLOGY OF THE CELL

fourth edition

Bruce Alberts

Alexander Johnson

Julian Lewis

Martin Raff

Keith Roberts

Peter Walter

GS Garland Science
Taylor & Francis Group

Garland
Vice President: Denise Schanck
Managing Editor: Sarah Gibbs
Senior Editorial Assistant: Kirsten Jenner
Managing Production Editor: Emma Hunt
Proofreader and Layout: Emma Hunt
Production Assistant: Angela Bennett
Text Editors: Marjorie Singer Anderson and Betsy Dilernia
Copy Editor: Bruce Goatly
Word Processors: Fran Dependahl, Misty Landers and Carol Winter
Designer: Blink Studio, London
Illustrator: Nigel Orme
Indexer: Janine Ross and Sherry Granum
Manufacturing: Nigel Eyre and Marion Morrow

Cell Biology Interactive
Artistic and Scientific Direction: Peter Walter
Narrated by: Julie Theriot
Production, Design, and Development: Mike Morales

Bruce Alberts received his Ph.D. from Harvard University and is President of the National Academy of Sciences and Professor of Biochemistry and Biophysics at the University of California, San Francisco. **Alexander Johnson** received his Ph.D. from Harvard University and is a Professor of Microbiology and Immunology at the University of California, San Francisco. **Julian Lewis** received his D.Phil. from the University of Oxford and is a Principal Scientist at the Imperial Cancer Research Fund, London. **Martin Raff** received his M.D. from McGill University and is at the Medical Research Council Laboratory for Molecular Cell Biology and Cell Biology Unit and in the Biology Department at University College London. **Keith Roberts** received his Ph.D. from the University of Cambridge and is Associate Research Director at the John Innes Centre, Norwich. **Peter Walter** received his Ph.D. from The Rockefeller University in New York and is Professor and Chairman of the Department of Biochemistry and Biophysics at the University of California, San Francisco, and an Investigator of the Howard Hughes Medical Institute.

General Library System
University of Wisconsin - Madison
728 State Street
Madison, WI 53706-1494
U.S.A.

© 2002 by Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter.
© 1983, 1989, 1994 by Bruce Alberts, Dennis Bray, Julian Lewis, Martin Raff, Keith Roberts, and James D. Watson.

All rights reserved. No part of this book covered by the copyright hereon may be reproduced or used in any format in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems—without permission of the publisher.

Library of Congress Cataloging-in-Publicaton Data
Molecular biology of the cell / Bruce Alberts ... [et al.].-- 4th ed.
p. cm
Includes bibliographical references and index.
ISBN 0-8153-3218-1 (hardbound) -- ISBN 0-8153-4072-9 (pbk.)
1. Cytology. 2. Molecular biology. I. Alberts, Bruce.
[DNLM: 1. Cells. 2. Molecular Biology.]
QH581.2 .M64 2002
571.6--dc21
2001054471 CIP

Published by Garland Science, a member of the Taylor & Francis Group, 29 West 35th Street, New York, NY 10001-2299

Printed in the United States of America

15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

Front cover Human Genome: Reprinted by permission from *Nature,* International Human Genome Sequencing Consortium, 409:860–921, 2001 © Macmillan Magazines Ltd. Adapted from an image by Francis Collins, NHGRI; Jim Kent, UCSC; Ewan Birney, EBI; and Darryl Leja, NHGRI; showing a portion of Chromosome 1 from the initial sequencing of the human genome.

Back cover In 1967, the British artist Peter Blake created a design classic. Nearly 35 years later Nigel Orme (illustrator), Richard Denyer (photographer), and the authors have together produced an affectionate tribute to Mr Blake's image. With its gallery of icons and influences, its assembly created almost as much complexity, intrigue and mystery as the original. *Drosophila, Arabidopsis,* Dolly and the assembled company tempt you to dip inside where, as in the original, "a splendid time is guaranteed for all."
(Gunter Blobel, courtesy of The Rockefeller University; Marie Curie, Keystone Press Agency Inc; Darwin bust, by permission of the President and Council of the Royal Society; Rosalind Franklin, courtesy of Cold Spring Harbor Laboratory Archives; Dorothy Hodgkin, © The Nobel Foundation, 1964; James Joyce, etching by Peter Blake; Robert Johnson, photo booth self-portrait early 1930s, © 1986 Delta Haze Corporation all rights reserved, used by permission; Albert L. Lehninger, (unidentified photographer) courtesy of The Alan Mason Chesney Medical Archives of The Johns Hopkins Medical Institutions; Linus Pauling, from Ava Helen and Linus Pauling Papers, Special Collections, Oregon State University; Nicholas Poussin, courtesy of ArtToday.com; Barbara McClintock, © David Micklos, 1983; Andrei Sakharov, courtesy of Elena Bonner; Frederick Sanger, © The Nobel Foundation, 1958.)

5454071

Preface

('611
QH
581.2
M64
2002
(201
rosw

"There is a paradox in the growth of scientific knowledge. As information accumulates in ever more intimidating quantities, disconnected facts and impenetrable mysteries give way to rational explanations, and simplicity emerges from chaos."

Thus began the preface of our first edition, written 18 years ago. Much of what we wrote in that preface holds for the present edition too. Our goals have not changed: we want to make cell biology comprehensible. We aim, as before, to give readers a perspective both on what is known, and on what is unknown. We have written the book for a wide range of students, but it should also prove useful for scientists wishing to follow progress outside their own specialized fields. As in past editions, each chapter is a joint composition of multiple authors and has been reviewed by a number of experts. This helps to explain why this edition, like the first, has been "a long time in gestation—three times longer than an elephant, five times longer than a whale."

Although information in the life sciences is expanding rapidly, human brain capacity is not. This discrepancy creates an increasing challenge for textbook writers, teachers, and especially students. We have been forced to think harder than ever in deciding what facts and concepts are essential. We have had to summarize—omitting much detail that, in overabundance, can prevent understanding. But specific examples are still needed to bring the subject to life. Thus, the most difficult part of the revision has been deciding what to leave out, while the most rewarding has been the opportunity that new discoveries provide to strengthen the conceptual framework.

What is new in the 4th edition?

Genomics has given us a new perspective that has demanded a complete recasting and expansion of the material on molecular genetics (Chapter 1 and Chapters 4 through 8). These six chapters can now be used as a stand-alone text in molecular biology. As before, the book ends with a major section on Cells in Their Social Context, but we have added a new chapter on Pathogens, Infection, and Innate Immunity. This reflects the remarkable advances in the understanding of the cell biology of infection, as well as the renewed awareness that infectious disease remains one of the greatest unconquered dangers in our world.

As our book makes clear, the complete sequencing of the genomes of hundreds of organisms, from bacteria to humans, has revolutionized our understanding of living things and the relationships between them. At last we can see what is there: the set of genes and proteins is finite, and we can list them. But we also recognize that these components are combined for use in marvelously subtle and complex ways, even in the simplest of organisms. Therefore, the traditional explanatory cartoons that we show on nearly every page of the book generally represent only the primitive first step toward an explanation. These drawings cannot capture the enormous complexity of the networks of protein–protein interactions that are responsible for most intracellular processes, whose understanding will require new and more quantitative forms of analysis. Thus, we are no longer as confident as we were 18 years ago that simplicity will eventually

emerge from the complexity. The extreme sophistication of cellular mechanisms will challenge cell biologists throughout the new century, which is very good news for the many young scientists who will succeed us.

As never before, new imaging and computer technologies have changed the ways we can observe the inner workings of living cells. We have tried to capture some of the excitement of these observations in *Cell Biology Interactive,* a CD-ROM disk that is included with each book. It contains dozens of video clips, animations, molecular structures, and high-resolution micrographs, which complement the static material in individual book chapters. One cannot watch cells crawling, dividing, segregating their chromosomes, or rearranging their surface without feeling curious about the molecular mechanisms that underlie these processes. We hope that *Cell Biology Interactive* will stimulate this curiosity in students and thereby make the learning of cell biology easier and more rewarding. We also hope that instructors will use these visual resources in the classroom for the same purpose.

We are deeply indebted to the many people who have helped us with this revision. The experts who critically reviewed specific chapters are acknowledged separately on p. xxix. We are especially grateful to Julie Theriot, who largely wrote both Chapter 16 and Chapter 25; we have all benefited from her wisdom. We are also grateful to the other experts who made major contributions to individual chapters: Nancy Craig helped to revise part of Chapter 5; Maynard Olson drafted the section on genome evolution in Chapter 7; Peter Shaw helped to revise Chapter 9; David Morgan largely wrote Chapter 17; Lisa Satterwhite prepared initial drafts of Chapter 18; Robert Kypta largely revised Chapter 19; Cori Bargmann helped us to restructure Chapter 21; and Paul Edwards played a central part in the revision of Chapter 23. Karen Hopkin drafted large parts of Chapters 8 and 23. We are also very grateful to the many scientists who generously provided micrographs for the book and materials for the CD-ROM.

Finally we are indebted to the outstanding staff of Garland Publishing. Nigel Orme oversaw the production of the final artwork with remarkable skill and speed. Mike Morales put in endless hours tackling the diverse challenges of developing the CD-ROM. Emma Hunt skillfully and artfully set the entire text and figures into pages. Sarah Gibbs and Kirsten Jenner kept us organized. Adam Sendroff and Nasreen Arain connected us to our customers. Eleanor Lawrence prepared the Glossary, and Mary Purton collected many of the references. Through it all, Denise Schanck calmly directed the whole effort with great skill. Last, but not least, Tim Hunt and John Wilson once again created a masterful Problems Book that supplements the main text; they worked side-by-side with us and were a constant source of wise advice. Their clever problems illustrate how discoveries are made and provide a unique way of learning cell biology.

It goes without saying that our book could not have been written without the strong support of our families and the forbearance of our friends, colleagues, and students. For decades, they have had to put up with our absences at frequent and lengthy book meetings, where most of the writing was done. They have helped us in innumerable ways, and we are grateful to them all.

Contents

Special Features *ix*
List of Topics *xi*
Acknowledgments *xxix*
A Note to the Reader *xxxiii*

PART I INTRODUCTION TO THE CELL

1. Cells and Genomes 3
2. Cell Chemistry and Biosynthesis 47
3. Proteins 129

PART II BASIC GENETIC MECHANISMS

4. DNA and Chromosomes 191
5. DNA Replication, Repair, and Recombination 235
6. How Cells Read the Genome: From DNA to Protein 299
7. Control of Gene Expression 375

PART III METHODS

8. Manipulating Proteins, DNA, and RNA 469
9. Visualizing Cells 547

PART IV INTERNAL ORGANIZATION OF THE CELL

10. Membrane Structure 583
11. Membrane Transport of Small Molecules and the Electrical Properties of Membranes 615
12. Intracellular Compartments and Protein Sorting 659
13. Intracellular Vesicular Traffic 711
14. Energy Conversion: Mitochondria and Chloroplasts 767
15. Cell Communication 831
16. The Cytoskeleton 907
17. The Cell Cycle and Programmed Cell Death 983
18. The Mechanics of Cell Division 1027

PART V CELLS IN THEIR SOCIAL CONTEXT

19. Cell Junctions, Cell Adhesion, and the Extracellular Matrix 1065
20. Germ Cells and Fertilization 1127
21. Development of Multicellular Organisms 1157
22. Histology: The Lives and Deaths of Cells in Tissues 1259
23. Cancer 1313
24. The Adaptive Immune System 1363
25. Pathogens, Infection, and Innate Immunity 1423

Glossary *G–1*
Index *I–1*
Tables: The Genetic Code, Amino Acids *T–1*

List of Topics

Part I Introduction to the Cell

CHAPTER 1 CELLS AND GENOMES 3

THE UNIVERSAL FEATURES OF CELLS ON EARTH

All Cells Store Their Hereditary Information in the Same Linear Chemical Code (DNA)
All Cells Replicate Their Hereditary Information by Templated Polymerization
All Cells Transcribe Portions of Their Hereditary Information into the Same Intermediary Form (RNA)
All Cells Use Proteins as Catalysts
All Cells Translate RNA into Protein in the Same Way
The Fragment of Genetic Information Corresponding to One Protein Is One Gene
Life Requires Free Energy
All Cells Function as Biochemical Factories Dealing with the Same Basic Molecular Building Blocks
All Cells Are Enclosed in a Plasma Membrane Across Which Nutrients and Waste Materials Must Pass
A Living Cell Can Exist with Fewer Than 500 Genes
Summary

THE DIVERSITY OF GENOMES AND THE TREE OF LIFE

Cells Can Be Powered by a Variety of Free Energy Sources
Some Cells Fix Nitrogen and Carbon Dioxide for Others
The Greatest Biochemical Diversity Is Seen Among Procaryotic Cells
The Tree of Life Has Three Primary Branches: Bacteria, Archaea, and Eucaryotes
Some Genes Evolve Rapidly; Others Are Highly Conserved
Most Bacteria and Archaea Have 1000–4000 Genes
New Genes Are Generated from Preexisting Genes
Gene Duplications Give Rise to Families of Related Genes Within a Single Cell
Genes Can Be Transferred Between Organisms, Both in the Laboratory and in Nature
Horizontal Exchanges of Genetic Information Within a Species Are Brought About by Sex
The Function of a Gene Can Often Be Deduced from Its Sequence
More Than 200 Gene Families Are Common to All Three Primary Branches of the Tree of Life
Mutations Reveal the Functions of Genes
Molecular Biologists Have Focused a Spotlight on *E. coli*
Summary

GENETIC INFORMATION IN EUCARYOTES

Eucaryotic Cells May Have Originated as Predators
Eucaryotic Cells Evolved From a Symbiosis
Eucaryotes Have Hybrid Genomes
Eucaryotic Genomes Are Big
Eucaryotic Genomes Are Rich in Regulatory DNA
The Genome Defines the Program of Multicellular Development
Many Eucaryotes Live as Solitary Cells: the Protists
A Yeast Serves as a Minimal Model Eucaryote
The Expression Levels of All The Genes of An Organism Can Be Monitored Simultaneously
Arabidopsis Has Been Chosen Out of 300,000 Species As a Model Plant
The World of Animal Cells Is Represented By a Worm, a Fly, a Mouse, and a Human
Studies in *Drosophila* Provide a Key to Vertebrate Development
The Vertebrate Genome Is a Product of Repeated Duplication
Genetic Redundancy Is a Problem for Geneticists, But It Creates Opportunities for Evolving Organisms
The Mouse Serves as a Model for Mammals
Humans Report on Their Own Peculiarities
We Are All Different in Detail
Summary
References

CHAPTER 2 CELL CHEMISTRY AND BIOSYNTHESIS 47

THE CHEMICAL COMPONENTS OF A CELL

Cells Are Made From a Few Types of Atoms
The Outermost Electrons Determine How Atoms Interact
Ionic Bonds Form by the Gain and Loss of Electrons
Covalent Bonds Form by the Sharing of Electrons
There Are Different Types of Covalent Bonds
An Atom Often Behaves as if It Has a Fixed Radius
Water Is the Most Abundant Substance in Cells
Some Polar Molecules Form Acids and Bases in Water
Four Types of Noncovalent Interactions Help Bring Molecules Together in Cells
A Cell Is Formed From Carbon Compounds
Cells Contain Four Major Families of Small Organic Molecules
Sugars Provide an Energy Source for Cells and Are the Subunits of Polysaccharides
Fatty Acids Are Components of Cell Membranes
Amino Acids Are the Subunits of Proteins
Nucleotides Are the Subunits of DNA and RNA
The Chemistry of Cells Is Dominated by Macromolecules with Remarkable Properties
Noncovalent Bonds Specify Both the Precise Shape of a Macromolecule and Its Binding to Other Molecules
Summary

CATALYSIS AND THE USE OF ENERGY BY CELLS

Cell Metabolism Is Organized by Enzymes
Biological Order Is Made Possible by the Release of Heat Energy from Cells
Photosynthetic Organisms Use Sunlight to Synthesize Organic Molecules
Cells Obtain Energy by the Oxidation of Organic Molecules
Oxidation and Reduction Involve Electron Transfers
Enzymes Lower the Barriers That Block Chemical Reactions
How Enzymes Find Their Substrates: The Importance of Rapid Diffusion
The Free-Energy Change for a Reaction Determines Whether It Can Occur
The Concentration of Reactants Influences ΔG
For Sequential Reactions, ΔG° Values Are Additive
Activated Carrier Molecules Are Essential for Biosynthesis
The Formation of an Activated Carrier Is Coupled to an Energetically Favorable Reaction
ATP Is the Most Widely Used Activated Carrier Molecule
Energy Stored in ATP Is Often Harnessed to Join Two Molecules Together
NADH and NADPH Are Important Electron Carriers
There Are Many Other Activated Carrier Molecules in Cells

The Synthesis of Biological Polymers Requires an Energy Input
Summary

HOW CELLS OBTAIN ENERGY FROM FOOD

Food Molecules Are Broken Down in Three Stages to Produce ATP
Glycolysis Is a Central ATP-Producing Pathway
Fermentations Allow ATP to Be Produced in the Absence of Oxygen
Glycolysis Illustrates How Enzymes Couple Oxidation to Energy Storage
Sugars and Fats Are Both Degraded to Acetyl CoA in Mitochondria
The Citric Acid Cycle Generates NADH by Oxidizing Acetyl Groups to CO_2
Electron Transport Drives the Synthesis of the Majority of the ATP in Most Cells
Organisms Store Food Molecules in Special Reservoirs
Amino Acids and Nucleotides Are Part of the Nitrogen Cycle
Many Biosynthetic Pathways Begin with Glycolysis or the Citric Acid Cycle
Metabolism Is Organized and Regulated
Summary
References

CHAPTER 3 PROTEINS 129

THE SHAPE AND STRUCTURE OF PROTEINS

The Shape of a Protein Is Specified by Its Amino Acid Sequence
Proteins Fold into a Conformation of Lowest Energy
The α Helix and the β Sheet Are Common Folding Patterns
The Protein Domain Is a Fundamental Unit of Organization
Few of the Many Possible Polypeptide Chains Will Be Useful
Proteins Can Be Classified into Many Families
Proteins Can Adopt a Limited Number of Different Protein Folds
Sequence Homology Searches Can Identify Close Relatives
Computational Methods Allow Amino Acid Sequences to Be Threaded into Known Protein Folds
Some Protein Domains, Called Modules, Form Parts of Many Different Proteins
The Human Genome Encodes a Complex Set of Proteins, Revealing Much That Remains Unknown
Larger Protein Molecules Often Contain More Than One Polypeptide Chain
Some Proteins Form Long Helical Filaments
A Protein Molecule Can Have an Elongated, Fibrous Shape
Extracellular Proteins Are Often Stabilized by Covalent Cross-Linkages
Protein Molecules Often Serve as Subunits for the Assembly of Large Structures
Many Structures in Cells Are Capable of Self-Assembly
The Formation of Complex Biological Structures Is Often Aided by Assembly Factors
Summary

PROTEIN FUNCTION

All Proteins Bind to Other Molecules
The Details of a Protein's Conformation Determine Its Chemistry
Sequence Comparisons Between Protein Family Members Highlight Crucial Ligand Binding Sites
Proteins Bind to Other Proteins Through Several Types of Interfaces
The Binding Sites of Antibodies Are Especially Versatile
Binding Strength Is Measured by the Equilibrium Constant
Enzymes Are Powerful and Highly Specific Catalysts
Substrate Binding Is the First Step in Enzyme Catalysis
Enzymes Speed Reactions by Selectively Stabilizing Transition States
Enzymes Can Use Simultaneous Acid and Base Catalysis
Lysozyme Illustrates How an Enzyme Works
Tightly Bound Small Molecules Add Extra Functions to Proteins
Multienzyme Complexes Help to Increase the Rate of Cell Metabolism
The Catalytic Activities of Enzymes Are Regulated
Allosteric Enzymes Have Two or More Binding Sites That Interact
Two Ligands Whose Binding Sites Are Coupled Must Reciprocally Affect Each Other's Binding
Symmetric Protein Assemblies Produce Cooperative Allosteric Transitions
The Allosteric Transition in Aspartate Transcarbamoylase Is Understood in Atomic Detail
Many Changes in Proteins Are Driven by Phosphorylation
A Eucaryotic Cell Contains a Large Collection of Protein Kinases and Protein Phosphatases
The Regulation of Cdk and Src Protein Kinases Shows How a Protein Can Function as a Microchip
Proteins That Bind and Hydrolyze GTP Are Ubiquitous Cellular Regulators
Regulatory Proteins Control the Activity of GTP-Binding Proteins by Determining Whether GTP or GDP Is Bound
Large Protein Movements Can Be Generated From Small Ones
Motor Proteins Produce Large Movements in Cells
Membrane-Bound Transporters Harness Energy to Pump Molecules Through Membranes
Proteins Often Form Large Complexes That Function as Protein Machines
A Complex Network of Protein Interactions Underlies Cell Function
Summary
References

Part II Basic Genetic Mechanisms

CHAPTER 4 DNA AND CHROMOSOMES 191

THE STRUCTURE AND FUNCTION OF DNA

A DNA Molecule Consists of Two Complementary Chains of Nucleotides
The Structure of DNA Provides a Mechanism for Heredity
In Eucaryotes, DNA Is Enclosed in a Cell Nucleus
Summary

CHROMOSOMAL DNA AND ITS PACKAGING IN THE CHROMATIN FIBER

Eucaryotic DNA Is Packaged into a Set of Chromosomes
Chromosomes Contain Long Strings of Genes
The Nucleotide Sequence of the Human Genome Shows How Genes Are Arranged in Humans
Comparisons Between the DNAs of Related Organisms Distinguish Conserved and Nonconserved Regions of DNA Sequence
Chromosomes Exist in Different States Throughout the Life of a Cell
Each DNA Molecule That Forms a Linear Chromosome Must Contain a Centromere, Two Telomeres, and Replication Origins
DNA Molecules Are Highly Condensed in Chromosomes
Nucleosomes Are the Basic Unit of Eucaryotic Chromosome Structure
The Structure of the Nucleosome Core Particle Reveals How DNA Is Packaged
The Positioning of Nucleosomes on DNA Is Determined by Both DNA Flexibility and Other DNA-Bound Proteins
Nucleosomes Are Usually Packed Together into a Compact Chromatin Fiber

ATP-Driven Chromatin Remodeling Machines Change Nucleosome Structure
Covalent Modification of the Histone Tails Can Profoundly Affect Chromatin
Summary

THE GLOBAL STRUCTURE OF CHROMOSOMES

Lampbrush Chromosomes Contain Loops of Decondensed Chromatin
Drosophila Polytene Chromosomes Are Arranged in Alternating Bands and Interbands
Both Bands and Interbands in Polytene Chromosomes Contain Genes
Individual Polytene Chromosome Bands Can Unfold and Refold as a Unit
Heterochromatin Is Highly Organized and Usually Resistant to Gene Expression
The Ends of Chromosomes Have a Special Form of Heterochromatin
Centromeres Are Also Packaged into Heterochromatin
Heterochromatin May Provide a Defense Mechanism Against Mobile DNA Elements
Mitotic Chromosomes Are Formed from Chromatin in Its Most Condensed State
Each Mitotic Chromosome Contains a Characteristic Pattern of Very Large Domains
Individual Chromosomes Occupy Discrete Territories in an Interphase Nucleus
Summary
References

CHAPTER 5 DNA REPLICATION, REPAIR, AND RECOMBINATION 235

THE MAINTENANCE OF DNA SEQUENCES

Mutation Rates Are Extremely Low
Many Mutations in Proteins Are Deleterious and Are Eliminated by Natural Selection
Low Mutation Rates Are Necessary for Life as We Know It
Summary

DNA REPLICATION MECHANISMS

Base-Pairing Underlies DNA Replication and DNA Repair
The DNA Replication Fork Is Asymmetrical
The High Fidelity of DNA Replication Requires Several Proofreading Mechanisms
Only DNA Replication in the 5′-to-3′ Direction Allows Efficient Error Correction
A Special Nucleotide-Polymerizing Enzyme Synthesizes Short RNA Primer Molecules on the Lagging Strand
Special Proteins Help to Open Up the DNA Double Helix in Front of the Replication Fork
A Moving DNA Polymerase Molecule Stays Connected to the DNA by a Sliding Ring
The Proteins at a Replication Fork Cooperate to Form a Replication Machine
A Strand-Directed Mismatch Repair System Removes Replication Errors That Escape From the Replication Machine
DNA Topoisomerases Prevent DNA Tangling During Replication
DNA Replication Is Similar in Eucaryotes and Bacteria
Summary

THE INITIATION AND COMPLETION OF DNA REPLICATION IN CHROMOSOMES

DNA Synthesis Begins at Replication Origins
Bacterial Chromosomes Have a Single Origin of DNA Replication
Eucaryotic Chromosomes Contain Multiple Origins of Replication
In Eucaryotes DNA Replication Takes Place During Only One Part of the Cell Cycle
Different Regions on the Same Chromosome Replicate at Distinct Times in S Phase
Highly Condensed Chromatin Replicates Late, While Genes in Less Condensed Chromatin Tend to Replicate Early
Well-Defined DNA Sequences Serve as Replication Origins in a Simple Eucaryote, the Budding Yeast
A Large Multisubunit Complex Binds to Eucaryotic Origins of Replication
The Mammalian DNA Sequences That Specify the Initiation of Replication Have Been Difficult to Identify
New Nucleosomes Are Assembled Behind the Replication Fork
Telomerase Replicates the Ends of Chromosomes
Telomere Length Is Regulated by Cells and Organisms
Summary

DNA REPAIR

Without DNA Repair, Spontaneous DNA Damage Would Rapidly Change DNA Sequences
The DNA Double Helix Is Readily Repaired
DNA Damage Can Be Removed by More Than One Pathway
The Chemistry of the DNA Bases Facilitates Damage Detection
Double-Strand Breaks are Efficiently Repaired
Cells Can Produce DNA Repair Enzymes in Response to DNA Damage
DNA Damage Delays Progression of the Cell Cycle
Summary

GENERAL RECOMBINATION

General Recombination Is Guided by Base-Pairing Interactions Between Two Homologous DNA Molecules
Meiotic Recombination Is Initiated by Double-Strand DNA Breaks
DNA Hybridization Reactions Provide a Simple Model for the Base-Pairing Step in General Recombination
The RecA Protein and its Homologs Enable a DNA Single Strand to Pair with a Homologous Region of DNA Double Helix
There Are Multiple Homologs of the RecA Protein in Eucaryotes, Each Specialized for a Specific Function
General Recombination Often Involves a Holliday Junction
General Recombination Can Cause Gene Conversion
General Recombination Events Have Different Preferred Outcomes in Mitotic and Meiotic Cells
Mismatch Proofreading Prevents Promiscuous Recombination Between Two Poorly Matched DNA Sequences
Summary

SITE-SPECIFIC RECOMBINATION

Mobile Genetic Elements Can Move by Either Transpositional or Conservative Mechanisms
Transpositional Site-Specific Recombination Can Insert a DNA Element into Any DNA Sequence
DNA-Only Transposons Move by DNA Breakage and Joining Reactions
Some Viruses Use Transpositional Site-Specific Recombination to Move Themselves into Host Cell Chromosomes
Retroviral-like Retrotransposons Resemble Retroviruses, but Lack a Protein Coat
A Large Fraction of the Human Genome Is Composed of Nonretroviral Retrotransposons
Different Transposable Elements Predominate in Different Organisms
Genome Sequences Reveal the Approximate Times When Transposable Elements Have Moved
Conservative Site-Specific Recombination Can Reversibly Rearrange DNA
Conservative Site-Specific Recombination Can be Used to Turn Genes On or Off
Summary
References

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

CHAPTER 6 HOW CELLS READ THE GENOME: FROM DNA TO PROTEIN 299

FROM DNA TO RNA

Portions of DNA Sequence Are Transcribed into RNA
Transcription Produces RNA Complementary to One Strand of DNA
Cells Produce Several Types of RNA
Signals Encoded in DNA Tell RNA Polymerase Where to Start and Stop
Transcription Start and Stop Signals Are Heterogeneous in Nucleotide Sequence
Transcription Initiation in Eucaryotes Requires Many Proteins
RNA Polymerase II Requires General Transcription Factors
Polymerase II Also Requires Activator, Mediator, and Chromatin-Modifying Proteins
Transcription Elongation Produces Superhelical Tension in DNA
Transcription Elongation in Eucaryotes Is Tightly Coupled To RNA Processing
RNA Capping Is the First Modification of Eucaryotic Pre-mRNAs
RNA Splicing Removes Intron Sequences from Newly Transcribed Pre-mRNAs
Nucleotide Sequences Signal Where Splicing Occurs
RNA Splicing Is Performed by the Spliceosome
The Spliceosome Uses ATP Hydrolysis to Produce a Complex Series of RNA–RNA Rearrangements
Ordering Influences in the Pre-mRNA Help to Explain How the Proper Splice Sites Are Chosen
A Second Set of snRNPs Splice a Small Fraction of Intron Sequences
RNA Splicing Shows Remarkable Plasticity
Spliceosome-Catalyzed RNA Splicing Probably Evolved from Self-Splicing Mechanisms
RNA-Processing Enzymes Generate the 3′ End of Eucaryotic mRNAs
Mature Eucaryotic mRNAs Are Selectively Exported from the Nucleus
Many Noncoding RNAs Are Also Synthesized and Processed in the Nucleus
The Nucleolus Is a Ribosome-Producing Factory
The Nucleus Contains a Variety of Subnuclear Structures
Summary

FROM RNA TO PROTEIN

An mRNA Sequence Is Decoded in Sets of Three Nucleotides
tRNA Molecules Match Amino Acids to Codons in mRNA
tRNAs Are Covalently Modified Before They Exit from the Nucleus
Specific Enzymes Couple Each Amino Acid to Its Appropriate tRNA Molecule
Editing by RNA Synthetases Ensures Accuracy
Amino Acids Are Added to the C-terminal End of a Growing Polypeptide Chain
The RNA Message Is Decoded on Ribosomes
Elongation Factors Drive Translation Forward
The Ribosome Is a Ribozyme
Nucleotide Sequences in mRNA Signal Where to Start Protein Synthesis
Stop Codons Mark the End of Translation
Proteins Are Made on Polyribosomes
Quality-Control Mechanisms Operate at Many Stages of Translation
There Are Minor Variations in the Standard Genetic Code
Many Inhibitors of Procaryotic Protein Synthesis Are Useful as Antibiotics
A Protein Begins to Fold While It Is Still Being Synthesized
Molecular Chaperones Help Guide the Folding of Many Proteins
Exposed Hydrophobic Regions Provide Critical Signals for Protein Quality Control
The Proteasome Degrades a Substantial Fraction of the Newly Synthesized Proteins in Cells
An Elaborate Ubiquitin-Conjugating System Marks Proteins for Destruction
Many Proteins Are Controlled by Regulated Destruction
Abnormally Folded Proteins Can Aggregate to Cause Destructive Human Diseases
There Are Many Steps From DNA to Protein
Summary

THE RNA WORLD AND THE ORIGINS OF LIFE

Life Requires Autocatalysis
Polynucleotides Can Both Store Information and Catalyze Chemical Reactions
A Pre-RNA World Probably Predates the RNA World
Single-Stranded RNA Molecules Can Fold into Highly Elaborate Structures
Self-Replicating Molecules Undergo Natural Selection
How Did Protein Synthesis Evolve?
All Present-Day Cells Use DNA as Their Hereditary Material
Summary
References

CHAPTER 7 CONTROL OF GENE EXPRESSION 375

AN OVERVIEW OF GENE CONTROL

The Different Cell Types of a Multicellular Organism Contain the Same DNA
Different Cell Types Synthesize Different Sets of Proteins
A Cell Can Change the Expression of Its Genes in Response to External Signals
Gene Expression Can Be Regulated at Many of the Steps in the Pathway from DNA to RNA to Protein
Summary

DNA-BINDING MOTIFS IN GENE REGULATORY PROTEINS

Gene Regulatory Proteins Were Discovered Using Bacterial Genetics
The Outside of the DNA Helix Can Be Read by Proteins
The Geometry of the DNA Double Helix Depends on the Nucleotide Sequence
Short DNA Sequences Are Fundamental Components of Genetic Switches
Gene Regulatory Proteins Contain Structural Motifs That Can Read DNA Sequences
The Helix–Turn–Helix Motif Is One of the Simplest and Most Common DNA-Binding Motifs
Homeodomain Proteins Constitute a Special Class of Helix–Turn–Helix Proteins
There Are Several Types of DNA-Binding Zinc Finger Motifs
β sheets Can Also Recognize DNA
The Leucine Zipper Motif Mediates Both DNA Binding and Protein Dimerization
Heterodimerization Expands the Repertoire of DNA Sequences Recognized by Gene Regulatory Proteins
The Helix–Loop–Helix Motif Also Mediates Dimerization and DNA Binding
It Is Not Yet Possible to Accurately Predict the DNA Sequences Recognized by All Gene Regulatory Proteins
A Gel-Mobility Shift Assay Allows Sequence-Specific DNA-Binding Proteins to Be Detected Readily
DNA Affinity Chromatography Facilitates the Purification of Sequence-Specific DNA-Binding Proteins
The DNA Sequence Recognized by a Gene Regulatory Protein Can Be Determined
A Chromatin Immunoprecipitation Technique Identifies DNA Sites Occupied by Gene Regulatory Proteins in Living Cells
Summary

Motif Exhibit 1059, 483 of 653

HOW GENETIC SWITCHES WORK

The Tryptophan Repressor Is a Simple Switch That Turns Genes On and Off in Bacteria
Transcriptional Activators Turn Genes On
A Transcriptional Activator and a Transcriptional Repressor Control the *lac* Operon
Regulation of Transcription in Eucaryotic Cells Is Complex
Eucaryotic Gene Regulatory Proteins Control Gene Expression From a Distance
A Eucaryotic Gene Control Region Consists of a Promoter Plus Regulatory DNA Sequences
Eucaryotic Gene Activator Proteins Promote the Assembly of RNA Polymerase and the General Transcription Factors at the Startpoint of Transcription
Eucaryotic Gene Activator Proteins Modify Local Chromatin Structure
Gene Activator Proteins Work Synergistically
Eucaryotic Gene Repressor Proteins Can Inhibit Transcription in Various Ways
Eucaryotic Gene Regulatory Proteins Often Assemble into Complexes on DNA
Complex Genetic Switches That Regulate *Drosophila* Development Are Built Up from Smaller Modules
The *Drosophila eve* Gene Is Regulated by Combinatorial Controls
Complex Mammalian Gene Control Regions Are Also Constructed From Simple Regulatory Modules
Insulators Are DNA Sequences That Prevent Eucaryotic Gene Regulatory Proteins From Influencing Distant Genes
Bacteria Use Interchangeable RNA Polymerase Subunits to Help Regulate Gene Transcription
Gene Switches Have Gradually Evolved
Summary

THE MOLECULAR GENETIC MECHANISMS THAT CREATE SPECIALIZED CELL TYPES

DNA Rearrangements Mediate Phase Variation in Bacteria
A Set of Gene Regulatory Proteins Determines Cell Type in a Budding Yeast
Two Proteins That Repress Each Other's Synthesis Determine the Heritable State of Bacteriophage Lambda
Gene Regulatory Circuits Can Be Used to Make Memory Devices as Well as Oscillators
Circadian Clocks Are Based on Feedback Loops in Gene Regulation
The Expression of Different Genes Can Be Coordinated by a Single Protein
Expression of a Critical Gene Regulatory Protein Can Trigger Expression of a Whole Battery of Downstream Genes
Combinatorial Gene Control Creates Many Different Cell Types in Eucaryotes
The Formation of an Entire Organ Can Be Triggered by a Single Gene Regulatory Protein
Stable Patterns of Gene Expression Can Be Transmitted to Daughter Cells
Chromosome Wide Alterations in Chromatin Structure Can Be Inherited
The Pattern of DNA Methylation Can Be Inherited When Vertebrate Cells Divide
Vertebrates Use DNA Methylation to Lock Genes in a Silent State
Genomic Imprinting Requires DNA Methylation
CG-Rich Islands Are Associated with About 20,000 Genes in Mammals
Summary

POSTTRANSCRIPTIONAL CONTROLS

Transcription Attenuation Causes the Premature Termination of Some RNA Molecules
Alternative RNA Splicing Can Produce Different Forms of a Protein From the Same Gene
The Definition of a Gene Has Had to Be Modified Since the Discovery of Alternative RNA Splicing
Sex Determination in *Drosophila* Depends on a Regulated Series of RNA Splicing Events
A Change in the Site of RNA Transcript Cleavage and Poly-A Addition Can Change the C-terminus of a Protein
RNA Editing Can Change the Meaning of the RNA Message
RNA Transport From the Nucleus Can Be Regulated
Some mRNAs Are Localized to Specific Regions of the Cytoplasm
Proteins That Bind to the 5′ and 3′ Untranslated Regions of mRNAs Mediate Negative Translational Control
The Phosphorylation of an Initiation Factor Globally Regulates Protein Synthesis
Initiation at AUG Codons Upstream of the Translation Start Can Regulate Eucaryotic Translation Initiation
Internal Ribosome Entry Sites Provide Opportunities for Translation Control
Gene Expression Can Be Controlled by a Change In mRNA Stability
Cytoplasmic Poly-A Addition Can Regulate Translation
Nonsense-Mediated mRNA Decay Is Used as an mRNA Surveillance System in Eucaryotes
RNA Interference Is Used by Cells to Silence Gene Expression
Summary

HOW GENOMES EVOLVE

Genome Alterations Are Caused by Failures of the Normal Mechanisms for Copying and Maintaining DNA
The Genome Sequences of Two Species Differ in Proportion to the Length of Time That They Have Separately Evolved
The Chromosomes of Humans and Chimpanzees Are Very Similar
A Comparison of Human and Mouse Chromosomes Shows How the Large-Scale Structures of Genomes Diverge
It Is Difficult to Reconstruct the Structure of Ancient Genomes
Gene Duplication and Divergence Provide a Critical Source of Genetic Novelty During Evolution
Duplicated Genes Diverge
The Evolution of the Globin Gene Family Shows How DNA Duplications Contribute to the Evolution of Organisms
Genes Encoding New Proteins Can Be Created by the Recombination of Exons
Genome Sequences Have Left Scientists with Many Mysteries to Be Solved
Genetic Variation within a Species Provides a Fine-Scale View of Genome Evolution
Summary
References

Part III Methods

CHAPTER 8 MANIPULATING PROTEINS, DNA, AND RNA 469

ISOLATING CELLS AND GROWING THEM IN CULTURE

Cells Can Be Isolated from a Tissue Suspension and Separated into Different Types
Cells Can Be Grown in a Culture Dish
Serum-Free, Chemically Defined Media Permit Identification of Specific Growth Factors
Eucaryotic Cell Lines Are a Widely Used Source of Homogeneous Cells
Cells Can Be Fused Together to Form Hybrid Cells
Hybridoma Cell Lines Provide a Permanent Source of Monoclonal Antibodies
Summary

FRACTIONATION OF CELLS

Organelles and Macromolecules Can Be Separated by Ultracentrifugation

The Molecular Details of Complex Cellular Processes Can Be Deciphered in Cell-Free Systems
Proteins Can Be Separated by Chromatography
Affinity Chromatography Exploits the Specfic Binding Sites on Proteins
The Size and Subunit Composition of a Protein Can Be Determined by SDS Polyacrylamide-Gel Electrophoresis
More Than 1000 Proteins Can Be Resolved on a Single Gel by Two-Dimensional Polyacrylamide-Gel Electrophoresis
Selective Cleavage of a Protein Generates a Distinctive Set of Peptide Fragments
Mass Spectrometry Can Be Used to Sequence Peptide Fragments and Identify Proteins
Summary

ISOLATING, CLONING, AND SEQUENCING DNA

Large DNA Molecules Are Cut into Fragments by Restriction Nucleases
Gel Electrophoresis Separates DNA Molecules of Different Sizes
Purified DNA Molecules Can Be Specifically Labeled with Radioisotopes or Chemical Markers *in vitro*
Nucleic Acid Hybridization Reactions Provide a Sensitive Way of Detecting Specific Nucleotide Sequences
Northern and Southern Blotting Facilitate Hybridization with Electrophoretically Separated Nucleic Acid Molecules
Hybridization Techniques Locate Specific Nucleic Acid Sequences in Cells or on Chromosomes
Genes Can Be Cloned from a DNA Library
Two Types of DNA Libraries Serve Different Purposes
cDNA Clones Contain Uninterrupted Coding Sequences
Isolated DNA Fragments Can Be Rapidly Sequenced
Nucleotide Sequences Are Used to Predict the Amino Acid Sequences of Proteins
The Genomes of Many Organisms Have Been Fully Sequenced
Selected DNA Segments Can Be Cloned in a Test Tube by a Polymerase Chain Reaction
Cellular Proteins Can Be Made in Large Amounts Through the Use of Expression Vectors
Summary

ANALYZING PROTEIN STRUCTURE AND FUNCTION

The Diffraction of X-rays by Protein Crystals Can Reveal a Protein's Exact Structure
Molecular Structure Can Also Be Determined Using Nuclear Magnetic Resonance (NMR) Spectroscopy
Sequence Similarity Can Provide Clues About Protein Function
Fusion Proteins Can Be Used to Analyze Protein Function and to Track Proteins in Living Cells
Affinity Chromatography and Immunoprecipitation Allow Identification of Associated Proteins
Protein–Protein Interactions Can Be Identified by Use of the Two-Hybrid System
Phage Display Methods Also Detect Protein Interactions
Protein Interactions Can Be Monitored in Real Time Using Surface Plasmon Resonance
DNA Footprinting Reveals the Sites Where Proteins Bind on a DNA Molecule
Summary

STUDYING GENE EXPRESSION AND FUNCTION

The Classical Approach Begins with Random Mutagenesis
Genetic Screens Identify Mutants Deficient in Cellular Processes
A Complementation Test Reveals Whether Two Mutations Are in the Same or in Different Genes
Genes Can Be Located by Linkage Analysis
Searching for Homology Can Help Predict a Gene's Function
Reporter Genes Reveal When and Where A Gene Is Expressed
Microarrays Monitor the Expression of Thousands of Genes At Once
Targeted Mutations Can Reveal Gene Function
Cells and Animals Containing Mutated Genes Can Be Made to Order
The Normal Gene in a Cell Can Be Directly Replaced by An Engineered Mutant Gene in Bacteria and Some Lower Eucaryotes
Engineered Genes Can Be Used to Create Specific Dominant Negative Mutations in Diploid Organisms
Gain-of-Function Mutations Provide Clues to the Role Genes Play in a Cell or Organism
Genes Can Be Redesigned to Produce Proteins of Any Desired Sequence
Engineered Genes Can Be Easily Inserted into the Germ Line of Many Animals
Gene Targeting Makes It Possible to Produce Transgenic Mice That Are Missing Specific Genes
Transgenic Plants Are Important for Both Cell Biology and Agriculture
Large Collections of Tagged Knockouts Provide a Tool for Examining the Function of Every Gene in an Organism
Summary
References

CHAPTER 9 VISUALIZING CELLS 547

LOOKING AT THE STRUCTURE OF CELLS IN THE MICROSCOPE

The Light Microscope Can Resolve Details 0.2 μm Apart
Living Cells Are Seen Clearly in a Phase-Contrast or a Differential-Interference-Contrast Microscope
Images Can Be Enhanced and Analyzed by Electronic Techniques
Tissues Are Usually Fixed and Sectioned for Microscopy
Different Components of the Cell Can Be Selectively Stained
Specific Molecules Can Be Located in Cells by Fluorescence Microscopy
Antibodies Can Be Used to Detect Specific Molecules
Imaging of Complex Three-Dimensional Objects Is Possible with the Optical Microscope
The Confocal Microscope Produces Optical Sections by Excluding Out-of-Focus Light
The Electron Microscope Resolves the Fine Structure of the Cell
Biological Specimens Require Special Preparation for the Electron Microscope
Specific Macromolecules Can Be Localized by Immunogold Electron Microscopy
Images of Surfaces Can Be Obtained by Scanning Electron Microscopy
Metal Shadowing Allows Surface Features to Be Examined at High Resolution by Transmission Electron Microscopy
Freeze-Fracture and Freeze-Etch Electron Microscopy Provide Views of Surfaces Inside the Cell
Negative Staining and Cryoelectron Microscopy Allow Macromolecules to Be Viewed at High Resolution
Multiple Images Can Be Combined to Increase Resolution
Views from Different Directions Can Be Combined to Give Three-Dimensional Reconstructions
Summary

VISUALIZING MOLECULES IN LIVING CELLS

Rapidly Changing Intracellular Ion Concentrations Can Be Measured with Light-Emitting Indicators
There Are Several Ways of Introducing Membrane-Impermeant Molecules into Cells
The Light-Induced Activation of "Caged" Precursor Molecules Facilitates Studies of Intracellular Dynamics
Green Fluorescent Protein Can Be Used to Tag Individual Proteins in Living Cells and Organisms
Light Can Be Used to Manipulate Microscopic Objects as well as to Image Them
Molecules Can Be Labeled with Radioisotopes
Radioisotopes Are Used to Trace Molecules in Cells and Organisms
Summary
References

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

Part IV Internal Organization of the Cell

CHAPTER 10 MEMBRANE STRUCTURE 583

THE LIPID BILAYER

Membrane Lipids Are Amphipathic Molecules, Most of Which Spontaneously Form Bilayers
The Lipid Bilayer Is a Two-Dimensional Fluid
The Fluidity of a Lipid Bilayer Depends on Its Composition
The Plasma Membrane Contains Lipid Rafts That Are Enriched in Sphingolipids, Cholesterol, and Some Membrane Proteins
The Asymmetry of the Lipid Bilayer Is Functionally Important
Glycolipids Are Found on the Surface of All Plasma Membranes
Summary

MEMBRANE PROTEINS

Membrane Proteins Can Be Associated with the Lipid Bilayer in Various Ways
In Most Transmembrane Proteins the Polypeptide Chain Crosses the Lipid Bilayer in an α-Helical Conformation
Some β Barrels Form Large Transmembrane Channels
Many Membrane Proteins Are Glycosylated
Membrane Proteins Can Be Solubilized and Purified in Detergents
The Cytosolic Side of Plasma Membrane Proteins Can Be Studied in Red Blood Cell Ghosts
Spectrin Is a Cytoskeletal Protein Noncovalently Associated with the Cytosolic Side of the Red Blood Cell Membrane
Glycophorin Extends Through the Red Blood Cell Lipid Bilayer as a Single α Helix
Band 3 of the Red Blood Cell Is a Multipass Membrane Protein That Catalyzes the Coupled Transport of Anions
Bacteriorhodopsin Is a Proton Pump That Traverses the Lipid Bilayer as Seven α Helices
Membrane Proteins Often Function as Large Complexes
Many Membrane Proteins Diffuse in the Plane of the Membrane
Cells Can Confine Proteins and Lipids to Specific Domains Within a Membrane
The Cell Surface Is Coated with Sugar Residues
Summary
References

CHAPTER 11 MEMBRANE TRANSPORT OF SMALL MOLECULES AND THE ELECTRICAL PROPERTIES OF MEMBRANES 615

PRINCIPLES OF MEMBRANE TRANSPORT

Protein-Free Lipid Bilayers Are Highly Impermeable to Ions
There Are Two Main Classes of Membrane Transport Proteins: Carriers and Channels
Active Transport Is Mediated by Carrier Proteins Coupled to an Energy Source
Ionophores Can Be Used as Tools to Increase the Permeability of Membranes to Specific Ions
Summary

CARRIER PROTEINS AND ACTIVE MEMBRANE TRANSPORT

Active Transport Can Be Driven by Ion Gradients
Na^+-Driven Carrier Proteins in the Plasma Membrane Regulate Cytosolic pH
An Asymmetric Distribution of Carrier Proteins in Epithelial Cells Underlies the Transcellular Transport of Solutes
The Plasma Membrane Na^+-K^+ Pump Is an ATPase
Some Ca^{2+} and H^+ Pumps Are Also P-type Transport ATPases
The Na^+-K^+ Pump Is Required to Maintain Osmotic Balance and Stabilize Cell Volume
Membrane-Bound Enzymes That Synthesize ATP Are Transport ATPases Working in Reverse
ABC Transporters Constitute the Largest Family of Membrane Transport Proteins
Summary

ION CHANNELS AND THE ELECTRICAL PROPERTIES OF MEMBRANES

Ion Channels Are Ion-Selective and Fluctuate Between Open and Closed States
The Membrane Potential in Animal Cells Depends Mainly on K^+ Leak Channels and the K^+ Gradient Across the Plasma Membrane
The Resting Potential Decays Only Slowly When the Na^+-K^+ Pump Is Stopped
The Three-Dimensional Structure of a Bacterial K^+ Channel Shows How an Ion Channel Can Work
The Function of a Nerve Cell Depends on Its Elongated Structure
Voltage-Gated Cation Channels Generate Action Potentials in Electrically Excitable Cells
Myelination Increases the Speed and Efficiency of Action Potential Propagation in Nerve Cells
Patch-Clamp Recording Indicates That Individual Gated Channels Open in an All-or-Nothing Fashion
Voltage-Gated Cation Channels Are Evolutionarily and Structurally Related
Transmitter-Gated Ion Channels Convert Chemical Signals into Electrical Ones at Chemical Synapses
Chemical Synapses Can Be Excitatory or Inhibitory
The Acetylcholine Receptors at the Neuromuscular Junction Are Transmitter-Gated Cation Channels
Transmitter-Gated Ion Channels Are Major Targets for Psychoactive Drugs
Neuromuscular Transmission Involves the Sequential Activation of Five Different Sets of Ion Channels
Single Neurons Are Complex Computation Devices
Neuronal Computation Requires a Combination of at Least Three Kinds of K^+ Channels
Long-term Potentiation (LTP) in the Mammalian Hippocampus Depends on Ca^{2+} Entry Through NMDA-Receptor Channels
Summary
References

Motif Exhibit 1059, 486 of 653

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

CHAPTER 12 INTRACELLULAR COMPARTMENTS AND PROTEIN SORTING 659

THE COMPARTMENTALIZATION OF CELLS
All Eucaryotic Cells Have the Same Basic Set of Membrane-Enclosed Organelles
The Topological Relationships of Membrane-Enclosed Organelles Can Be Interpreted in Terms of Their Evolutionary Origins
Proteins Can Move Between Compartments in Different Ways
Signal Sequences and Signal Patches Direct Proteins to the Correct Cellular Address
Most Membrane-Enclosed Organelles Cannot Be Constructed From Scratch: They Require Information in the Organelle Itself
Summary

THE TRANSPORT OF MOLECULES BETWEEN THE NUCLEUS AND THE CYTOSOL
Nuclear Pore Complexes Perforate the Nuclear Envelope
Nuclear Localization Signals Direct Nuclear Proteins to the Nucleus
Nuclear Import Receptors Bind Nuclear Localization Signals and Nucleoporins
Nuclear Export Works Like Nuclear Import, But in Reverse
The Ran GTPase Drives Directional Transport Through Nuclear Pore Complexes
Transport Between the Nucleus and Cytosol Can Be Regulated by Controlling Access to the Transport Machinery
The Nuclear Envelope Is Disassembled During Mitosis
Summary

THE TRANSPORT OF PROTEINS INTO MITOCHONDRIA AND CHLOROPLASTS
Translocation into the Mitochondrial Matrix Depends on a Signal Sequence and Protein Translocators
Mitochondrial Precursor Proteins Are Imported as Unfolded Polypeptide Chains
Mitochondrial Precursor Proteins Are Imported into the Matrix at Contact Sites That Join the Inner and Outer Membranes
ATP Hydrolysis and a H^+ Gradient are Used to Drive Protein Import into Mitochondria
Repeated Cycles of ATP Hydrolysis by Mitochondrial Hsp70 Complete the Import Process
Protein Transport into the Inner Mitochondrial Membrane and the Intermembrane Space Requires Two Signal Sequences
Two Signal Sequences Are Required to Direct Proteins to the Thylakoid Membrane in Chloroplasts
Summary

PEROXISOMES
Peroxisomes Use Molecular Oxygen and Hydrogen Peroxide to Perform Oxidative Reactions
A Short Signal Sequence Directs the Import of Proteins into Peroxisomes
Summary

THE ENDOPLASMIC RETICULUM
Membrane-Bound Ribosomes Define the Rough ER
Smooth ER Is Abundant in Some Specialized Cells
Rough and Smooth Regions of ER Can Be Separated by Centrifugation
Signal Sequences Were First Discovered in Proteins Imported into the Rough ER
A Signal-Recognition Particle (SRP) Directs ER Signal Sequences to a Specific Receptor in the Rough ER Membrane
The Polypeptide Chain Passes Through an Aqueous Pore in the Translocator
Translocation Across the ER Membrane Does Not Always Require Ongoing Polypeptide Chain Elongation
The ER Signal Sequnce Is Removed from Most Soluble Proteins After Translocation
In Single-Pass Transmembrane Proteins, a Single Internal ER Signal Sequence Remains in the Lipid Bilayer as a Membrane-Spanning α helix
Combinations of Start-Transfer and Stop-Transfer Signals Determine the Topology of Multipass Transmembrane Proteins
Translocated Polypeptide Chains Fold and Assemble in the Lumen of the Rough ER
Most Proteins Synthesized in the Rough ER Are Glycosylated by the Addition of a Common *N*-Linked Oligosaccharide
Oligosaccharides Are Used as Tags to Mark the State of Protein Folding
Improperly Folded Proteins Are Exported from the ER and Degraded in the Cytosol
Misfolded Proteins in the ER Activate an Unfolded Protein Response
Some Membrane Proteins Acquire A Covalently Attached Glycosylphosphatidylinositol (GPI) Anchor
Most Membrane Lipid Bilayers Are Assembled in the ER
Phospholipid Exchange Proteins Help to Transport Phospholipids from the ER to Mitochondria and Peroxisomes
Summary
References

CHAPTER 13 INTRACELLULAR VESICULAR TRAFFIC 711

THE MOLECULAR MECHANISMS OF MEMBRANE TRANSPORT AND THE MAINTENANCE OF COMPARTMENTAL DIVERSITY
There Are Various Types of Coated Vesicles
The Assembly of a Clathrin Coat Drives Vesicle Formation
Both The Pinching-Off and Uncoating of Coated Vesicles Are Regulated Processes
Not All Transport Vesicles are Spherical
Monomeric GTPases Control Coat Assembly
SNARE Proteins and Targeting GTPases Guide Membrane Transport
Interacting SNAREs Need To Be Pried Apart Before They Can Function Again
Rab Proteins Help Ensure the Specificity of Vesicle Docking
SNAREs May Mediate Membrane Fusion
Viral Fusion Proteins and SNAREs May Use Similar Strategies
Summary

TRANSPORT FROM THE ER THROUGH THE GOLGI APPARATUS
Proteins Leave the ER in COPII-Coated Transport Vesicles
Only Proteins That Are Properly Folded and Assembled Can Leave the ER
Transport from the ER to the Golgi Apparatus Is Mediated by Vesicular Tubular Clusters
The Retrieval Pathway to the ER Uses Sorting Signals
Many Proteins are Selectively Retained in the Compartments in which they Function
The Length of the Transmembrane Region of Golgi Enzymes Determines their Location in The Cell
The Golgi Apparatus Consists of an Ordered Series of Compartments
Oligosaccharide Chains Are Processed in the Golgi Apparatus
Proteoglycans Are Assembled in the Golgi Apparatus
What Is the Purpose of Glycosylation?

Motif Exhibit 1059, 487 of 653

The Golgi Cisternae Are Organized as a Series of Processing Compartments
Transport Through the Golgi Apparatus May Occur by Vesicular Transport or Cisternal Maturation
Matrix Proteins Form a Dynamic Scaffold That Helps Organize the Apparatus
Summary

TRANSPORT FROM THE *TRANS* GOLGI NETWORK TO LYSOSOMES

Lysosomes Are the Principal Sites of Intracellular Digestion
Lysosomes Are Heterogeneous
Plant and Fungal Vacuoles Are Remarkably Versatile Lysosomes
Multiple Pathways Deliver Materials to Lysosomes
A Mannose 6-Phosphate Receptor Recognizes Lysosomal Proteins in the *Trans* Golgi Network
The M6P Receptor Shuttles Between Specific Membranes
A Signal Patch in the Hydrolyase Polypeptide Chain Provides the Cue for M6P Addition
Defects in the GlcNAc Phosphotransferase Cause a Lysosomal Storage Disease in Humans
Some Lysosomes May Undergo Exocytosis
Summary

TRANSPORT INTO THE CELL FROM THE PLASMA MEMBRANE: ENDOCYTOSIS

Specialized Phagocytic Cells Can Ingest Large Particles
Pinocytic Vesicles Form from Coated Pits in the Plasma Membrane
Not All Pinocytic Vesicles Are Clathrin-Coated
Cells Import Selected Extracellular Macromolecules by Receptor-Mediated Endocytosis
Endocytosed Materials That Are Not Retrieved From Endosomes End Up in Lysosomes
Specific Proteins Are Removed From Early Endosomes and Returned to the Plasma Membrane
Multivesicular Bodies Form on the Pathway to Late Endosomes
Macromolecules Can Be Transferred Across Epithelial Cell Sheets by Transcytosis
Epithelial Cells Have Two Distinct Early Endosomal Compartments But a Common Late Endosomal Compartment
Summary

TRANSPORT FROM THE *TRANS* GOLGI NETWORK TO THE CELL EXTERIOR: EXOCYTOSIS

Many Proteins and Lipids Seem to Be Carried Automatically From the Golgi Apparatus to the Cell Surface
Secretory Vesicles Bud From the *Trans* Golgi Network
Proteins Are Often Proteolytically Processed During the Formation of Secretory Vesicles
Secretory Vesicles Wait Near the Plasma Membrane Until Signaled to Release Their Contents
Regulated Exocytosis Can Be a Localized Response of the Plasma Membrane and Its Underlying Cytoplasm
Secretory Vesicle Membrane Components Are Quickly Removed From the Plasma Membrane
Polarized Cells Direct Proteins From the *Trans* Golgi Network to the Appropriate Domain of the Plasma Membrane
Cytoplasmic Sorting Signals Guide Membrane Proteins Selectively to the Basolateral Plasma Membrane
Lipid Rafts May Mediate Sorting of Glycosphingolipids and GPI-Anchored Proteins to the Apical Plasma Membrane
Synaptic Vesicles Can Form Directly from Endocytic Vesicles
Summary
References

CHAPTER 14 ENERGY CONVERSION: MITOCHONDRIA AND CHLOROPLASTS 767

THE MITOCHONDRION

The Mitochondrion Contains an Outer Membrane, an Inner Membrane, and Two Internal Compartments
High-Energy Electrons Are Generated via the Citric Acid Cycle
A Chemiosmotic Process Converts Oxidation Energy into ATP
Electrons Are Transferred From NADH to Oxygen Through Three Large Respiratory Enzyme Complexes
As Electrons Move Along the Respiratory Chain, Energy Is Stored as an Electrochemical Proton Gradient Across the Inner Membrane
How the Proton Gradient Drives ATP Synthesis
How the Proton Gradient Drives Coupled Transport Across the Inner Membrane
Proton Gradients Produce Most of the Cell's ATP
Mitochondria Maintain a High ATP:ADP Ratio in Cells
A Large Negative Value of ΔG for ATP Hydrolysis Makes ATP Useful to the Cell
ATP Synthase Can Also Function in Reverse to Hydrolyze ATP and Pump H^+
Summary

ELECTRON-TRANSPORT CHAINS AND THEIR PROTON PUMPS

Protons Are Unusually Easy to Move
The Redox Potential Is a Measure of Electron Affinities
Electron Transfers Release Large Amounts of Energy
Spectroscopic Methods Have Been Used to Identify Many Electron Carriers in the Respiratory Chain
The Respiratory Chain Includes Three Large Enzyme Complexes Embedded in the Inner Membrane
An Iron–Copper Center in Cytochrome Oxidase Catalyzes Efficient O_2 Reduction
Electron Transfers Are Mediated by Random Collisions in the Inner Mitochondrial Membrane
A Large Drop in Redox Potential Across Each of the Three Respiratory Enzyme Complexes Provides the Energy for H^+ Pumping
The Mechanism of H^+ Pumping Will Soon Be Understood in Atomic Detail
H^+ Ionophores Uncouple Electron Transport From ATP Synthesis
Respiratory Control Normally Restrains Electron Flow Through the Chain
Natural Uncouplers Convert the Mitochondria in Brown Fat into Heat-Generating Machines
Bacteria Also Exploit Chemiosmotic Mechanisms to Harness Energy
Summary

CHLOROPLASTS AND PHOTOSYNTHESIS

The Chloroplast Is One Member of the Plastid Family of Organelles
Chloroplasts Resemble Mitochondria But Have an Extra Compartment
Chloroplasts Capture Energy From Sunlight and Use It to Fix Carbon
Carbon Fixation Is Catalyzed by Ribulose Bisphosphate Carboxylase
Three Molecules of ATP and Two Molecules of NADPH Are Consumed for Each CO_2 Molecule That Is Fixed
Carbon Fixation in Some Plants Is Compartmentalized to Facilitate Growth at Low CO_2 Concentrations
Photosynthesis Depends on the Photochemistry of Chlorophyll Molecules
A Photosystem Consists of a Reaction Center Plus an Antenna Complex
In a Reaction Center, Light Energy Captured by Chlorophyll Creates a Strong Electron Donor from a Weak One
Noncyclic Photophosphorylation Produces Both NADPH and ATP
Chloroplasts Can Make ATP by Cyclic Photophosphorylation Without Making NADPH

Photosystems I and II Have Related Structures, and Also Resemble Bacterial Photosystems
The Proton-Motive Force Is the Same in Mitochondria and Chloroplasts
Carrier Proteins in the Chloroplast Inner Membrane Control Metabolite Exchange with the Cytosol
Chloroplasts Also Perform Other Crucial Biosyntheses
Summary

THE GENETIC SYSTEMS OF MITOCHONDRIA AND PLASTIDS

Mitochondria and Chloroplasts Contain Complete Genetic Systems
Organelle Growth and Division Determine the Number of Mitochondria and Plastids in a Cell
The Genomes of Mitochondria and Chloroplasts Are Diverse
Mitochondria and Chloroplasts Probably Both Evolved from Endosymbiotic Bacteria
Mitochondrial Genomes Have Several Surprising Features
Animal Mitochondria Contain the Simplest Genetic Systems Known
Some Organelle Genes Contain Introns
The Chloroplast Genome of Higher Plants Contains About 120 Genes
Mitochondrial Genes Are Inherited by a Non-Mendelian Mechanism
Organelle Genes Are Maternally Inherited in Many Organisms
Petite Mutants in Yeasts Demonstrate the Overwhelming Importance of the Cell Nucleus for Mitochondrial Biogenesis
Mitochondria and Plastids Contain Tissue-Specific Proteins that Are Encoded in the Cell Nucleus
Mitochondria Import Most of Their Lipids; Chloroplasts Make Most of Theirs
Why Do Mitochondria and Chloroplasts Have Their Own Genetic Systems?
Summary

THE EVOLUTION OF ELECTRON-TRANSPORT CHAINS

The Earliest Cells Probably Produced ATP by Fermentation
Electron-transport Chains Enabled Anaerobic Bacteria to Use Nonfermentable Molecules as Their Major Source of Energy
By Providing an Inexhaustible Source of Reducing Power, Photosynthetic Bacteria Overcame a Major Evolutionary Obstacle
The Photosynthetic Electron-Transport Chains of Cyanobacteria Produced Atmospheric Oxygen and Permitted New Life-Forms
Summary
References

CHAPTER 15 CELL COMMUNICATION 831

GENERAL PRINCIPLES OF CELL COMMUNICATION

Extracellular Signal Molecules Bind to Specific Receptors
Extracellular Signal Molecules Can Act Over Either Short or Long Distances
Autocrine Signaling Can Coordinate Decisions by Groups of Identical Cells
Gap Junctions Allow Signaling Information to Be Shared by Neighboring Cells
Each Cell Is Programmed to Respond to Specific Combinations of Extracellular Signal Molecules
Different Cells Can Respond Differently to the Same Extracellular Signal Molecule
The Concentration of a Molecule Can Be Adjusted Quickly Only If the Lifetime of the Molecule Is Short
Nitric Oxide Gas Signals by Binding Directly to an Enzyme Inside the Target Cell
Nuclear Receptors Are Ligand-Activated Gene Regulatory Proteins
The Three Largest Classes of Cell-Surface Receptor Proteins Are Ion-Channel-Linked, G-Protein-Linked, and Enzyme-Linked Receptors
Most Activated Cell-Surface Receptors Relay Signals Via Small Molecules and a Network of Intracellular Signaling Proteins
Some Intracellular Signaling Proteins Act as Molecular Switches
Intracellular Signaling Complexes Enhance the Speed, Efficiency, and Specificity of the Response
Interactions Between Intracellular Signaling Proteins Are Mediated by Modular Binding Domains
Cells Can Respond Abruptly to a Gradually Increasing Concentration of an Extracellular Signal
A Cell Can Remember The Effect of Some Signals
Cells Can Adjust Their Sensitivity to a Signal
Summary

SIGNALING THROUGH G-PROTEIN-LINKED CELL-SURFACE RECEPTORS

Trimeric G Proteins Disassemble to Relay Signals from G-Protein-Linked Receptors
Some G Proteins Signal By Regulating the Production of Cyclic AMP
Cyclic-AMP-Dependent Protein Kinase (PKA) Mediates Most of the Effects of Cyclic AMP
Protein Phosphatases Make the Effects of PKA and Other Protein Kinases Transitory
Some G Proteins Activate the Inositol Phospholipid Signaling Pathway by Activating Phospholipase C-β
Ca^{2+} Functions as a Ubiquitous Intracellular Messenger
The Frequency of Ca^{2+} Oscillations Influences a Cell's Response
Ca^{2+}/Calmodulin-Dependent Protein Kinases (CaM-Kinases) Mediate Many of the Actions of Ca^{2+} in Animal Cells
Some G Proteins Directly Regulate Ion Channels
Smell and Vision Depend on G-Protein-Linked Receptors That Regulate Cyclic-Nucleotide-Gated Ion Channels
Extracellular Signals Are Greatly Amplified by the Use of Small Intracellular Mediators and Enzymatic Cascades
G-Protein-Linked Receptor Desensitization Depends on Receptor Phosphorylation
Summary

SIGNALING THROUGH ENZYME-LINKED CELL-SURFACE RECEPTORS

Activated Receptor Tyrosine Kinases Phosphorylate Themselves
Phosphorylated Tyrosines Serve as Docking Sites For Proteins With SH2 Domains
Ras Is Activated by a Guanine Nucleotide Exchange Factor
Ras Activates a Downstream Serine/Threonine Phosphorylation Cascade That Includes a MAP-Kinase
PI 3-Kinase Produces Inositol Phospholipid Docking Sites in the Plasma Membrane
The PI 3-Kinase/Protein Kinase B Signaling Pathway Can Stimulate Cells to Survive and Grow
Tyrosine-Kinase-Associated Receptors Depend on Cytoplasmic Tyrosine Kinases for Their Activity
Cytokine Receptors Activate the Jak–STAT Signaling Pathway, Providing a Fast Track to the Nucleus
Some Protein Tyrosine Phosphatases May Act as Cell-Surface Receptors
Signal Proteins of the TGF-β Superfamily Act Through Receptor Serine/Threonine Kinases and Smads
Receptor Guanylyl Cyclases Generate Cyclic GMP Directly
Bacterial Chemotaxis Depends on a Two-Component Signaling Pathway Activated by Histidine–Kinase-Associated Receptors
Summary

SIGNALING PATHWAYS THAT DEPEND ON REGULATED PROTEOLYSIS

The Receptor Protein Notch Is Activated by Cleavage
Wnt Proteins Bind to Frizzled Receptors and Inhibit the Degradation of β-Catenin
Hedgehog Proteins Act Through a Receptor Complex of Patched and Smoothened, Which Oppose Each Other
Multiple Stressful and Proinflammatory Stimuli Act Through an NF-κB-Dependent Signaling Pathway
Summary

SIGNALING IN PLANTS

Multicellularity and Cell Communication Evolved Independently in Plants and Animals
Receptor Serine/Threonine Kinases Function as Cell-Surface Receptors in Plants
Ethylene Activates a Two-component Signaling Pathway
Phytochromes Detect Red Light, and Cryptochromes Detect Blue Light
Summary
References

CHAPTER 16 THE CYTOSKELETON 907

THE SELF-ASSEMBLY AND DYNAMIC STRUCTURE OF CYTOSKELETAL FILAMENTS

Each Type of Cytoskeletal Filament Is Constructed from Smaller Protein Subunits
Filaments Formed from Multiple Protofilaments Have Advantageous Properties
Nucleation Is the Rate-Limiting Step in the Formation of a Cytoskeletal Polymer
The Tubulin and Actin Subunits Assemble Head-to-Tail, Creating Filaments that Are Polar
The Two Ends of a Microtubule and of an Actin Filament Are Distinct and Grow at Different Rates
Filament Treadmilling and Dynamic Instability Are Consequences of Nucleotide Hydrolysis by Tubulin and Actin
Treadmilling and Dynamic Instability Require Energy but Are Useful
Other Polymeric Proteins Also Use Nucleotide Hydrolysis to Couple a Conformational Change to Cell Movements
Tubulin and Actin Have Been Highly Conserved During Eucaryotic Evolution
Intermediate Filament Structure Depends on The Lateral Bundling and Twisting of Coiled Coils
Intermediate Filaments Impart Mechanical Stability to Animal Cells
Filament Polymerization Can Be Altered by Drugs
Summary

HOW CELLS REGULATE THEIR CYTOSKELETAL FILAMENTS

Microtubules Are Nucleated by a Protein Complex Containing γ-tubulin
Microtubules Emanate from the Centrosome in Animal Cells
Actin Filaments Are Often Nucleated at the Plasma Membrane
Filament Elongation Is Modified by Proteins That Bind to the Free Subunits
Proteins That Bind Along the Sides of Filaments Can Either Stabilize or Destabilize Them
Proteins That Interact with Filament Ends Can Dramatically Change Filament Dynamics
Filaments Are Organized into Higher-Order Structures in Cells
Intermediate Filaments Are Cross-Linked and Bundled Into Strong Arrays
Cross-Linking Proteins with Distinct Properties Organize Different Actin Assemblies
Severing Proteins Regulate the Length and Kinetic Behavior of Actin Filaments and Microtubules
Cytoskeletal Elements Can Attach to the Plasma Membrane
Special Bundles of Cytoskeletal Filaments Form Strong Attachments Across the Plasma Membrane: Focal Contacts, Adhesion Belts, and Desmosomes
Extracellular Signals Can Induce Major Cytoskeletal Rearrangements
Summary

MOLECULAR MOTORS

Actin-Based Motor Proteins Are Members of the Myosin Superfamily
There Are Two Types of Microtubule Motor Proteins: Kinesins and Dyneins
The Structural Similarity of Myosin and Kinesin Indicates a Common Evolutionary Origin
Motor Proteins Generate Force by Coupling ATP Hydrolysis to Conformational Changes
Motor Protein Kinetics Are Adapted to Cell Functions
Motor Proteins Mediate the Intracellular Transport of Membrane-Enclosed Organelles
Motor Protein Function Can Be Regulated
Muscle Contraction Depends on the Sliding of Myosin II and Actin Filaments
Muscle Contraction Is Initiated by a Sudden Rise in Cytosolic Ca^{2+} Concentration
Heart Muscle Is a Precisely Engineered Machine
Cilia and Flagella Are Motile Structures Built from Microtubules and Dyneins
Summary

THE CYTOSKELETON AND CELL BEHAVIOR

Mechanisms of Cell Polarization Can Be Readily Analyzed in Yeast Cells
Specific RNA Molecules Are Localized by the Cytoskeleton
Many Cells Can Crawl Across A Solid Substratum
Plasma Membrane Protrusion Is Driven by Actin Polymerization
Cell Adhesion and Traction Allow Cells to Pull Themselves Forward
External Signals Can Dictate the Direction of Cell Migration
The Complex Morphological Specialization of Neurons Depends on The Cytoskeleton
Summary
References

CHAPTER 17 THE CELL CYCLE AND PROGRAMMED CELL DEATH 983

AN OVERVIEW OF THE CELL CYCLE

The Cell-Cycle Control System Is Similar in All Eucaryotes
The Cell-Cycle Control System Can Be Dissected Genetically in Yeasts
The Cell-Cycle Control System Can Be Analyzed Biochemically in Animal Embryos
The Cell-Cycle Control System in Mammals Can Be Studied in Culture
Cell-Cycle Progression Can Be Studied in Various Ways
Summary

COMPONENTS OF THE CELL-CYCLE CONTROL SYSTEM

The Cell-Cycle Control System Triggers the Major Processes of the Cell Cycle
The Control System Can Arrest the Cell Cycle at Specific Checkpoints
Checkpoints Generally Operate Through Negative Intracellular Signals
The Cell-Cycle Control System Is Based on Cyclically Activated Protein Kinases
Cdk Activity Can Be Suppressed Both by Inhibitory Phosphorylation and by Inhibitory Proteins
The Cell-Cycle Control System Depends on Cyclical Proteolysis
Cell-Cycle Control Also Depends on Transcriptional Regulation
Summary

INTRACELLULAR CONTROL OF CELL-CYCLE EVENTS

S-Phase Cyclin–Cdk Complexes (S-Cdks) Initiate DNA Replication Once Per Cycle
The Activation of M-Phase Cyclin–Cdk Complexes (M-Cdks) Triggers Entry into Mitosis

Entry into Mitosis Is Blocked by Incomplete DNA Replication: The DNA Replication Checkpoint
M-Cdk Prepares the Duplicated Chromosomes for Separation
Sister Chromatid Separation Is Triggered by Proteolysis
Unattached Chromosomes Block Sister-Chromatid Separation: The Spindle-Attachment Checkpoint
Exit from Mitosis Requires the Inactivation of M-Cdk
The G_1 Phase Is a State of Stable Cdk Inactivity
The Rb Protein Acts as a Brake in Mammalian G_1 Cells
Cell-Cycle Progression Is Somehow Co-ordinated With Cell Growth
Cell-Cycle Progression is Blocked by DNA Damage and p53: DNA Damage Checkpoints
Summary

PROGRAMMED CELL DEATH (APOPTOSIS)

Apoptosis Is Mediated by an Intracellular Proteolytic Cascade
Procaspases Are Activated by Binding to Adaptor Proteins
Bcl-2 Family Proteins and IAP Proteins Are the Main Intracellular Regulators of the Cell Death Program
Summary

EXTRACELLULAR CONTROL OF CELL DIVISION, CELL GROWTH, AND APOPTOSIS

Mitogens Stimulate Cell Division
Cells Can Delay Division by Entering a Specialized Nondividing State
Mitogens Stimulate G_1-Cdk and G_1/S-Cdk Activities
Abnormal Proliferation Signals Cause Cell-Cycle Arrest or Cell Death
Human Cells Have a Built-in Limitation on the Number of Times They Can Divide
Extracellular Growth Factors Stimulate Cell Growth
Extracellular Survival Factors Suppress Apoptosis
Neighboring Cells Compete for Extracellular Signal Proteins
Many Types of Normal Animal Cells Need Anchorage to Grow and Proliferate
Some Extracellular Signal Proteins Inhibit Cell Growth, Cell Division, and Survival
Intricately Regulated Patterns of Cell Division Generate and Maintain Body Form
Summary
References

CHAPTER 18 THE MECHANICS OF CELL DIVISION 1027

AN OVERVIEW OF M PHASE

Cohesins and Condensins Help Configure Replicated Chromosomes for Segregation
Cytoskeletal Machines Perform Both Mitosis and Cytokinesis
Two Mechanisms Help Ensure That Mitosis Always Precedes Cytokinesis
M Phase in Animal Cells Depends on Centrosome Duplication in the Preceding Interphase
M Phase Is Traditionally Divided into Six Stages
Summary

MITOSIS

Microtubule Instability Increases Greatly at M Phase
Interactions Between Opposing Motor Proteins and Microtubules of Opposite Polarity Drive Spindle Assembly
Kinetochores Attach Chromosomes to the Mitotic Spindle
Microtubules Are Highly Dynamic in the Metaphase Spindle
Functional Bipolar Spindles Can Assemble Around Chromosomes in Cells Without Centrosomes
Anaphase Is Delayed Until All Chromosomes Are Positioned at the Metaphase Plate
Sister Chromatids Separate Suddenly at Anaphase
Kinetochore Microtubules Disassemble at Both Ends During Anaphase A
Both Pushing and Pulling Forces Contribute to Anaphase B
At Telophase, the Nuclear Envelope Re-forms Around Individual Chromosomes
Summary

CYTOKINESIS

The Microtubules of the Mitotic Spindle Determine the Plane of Animal Cell Division
Some Cells Reposition Their Spindle to Divide Asymmetrically
Actin and Myosin II in the Contractile Ring Generate the Force for Cytokinesis
Membrane-Enclosed Organelles Must Be Distributed to Daughter Cells During Cytokinesis
Mitosis Can Occur Without Cytokinesis
The Phragmoplast Guides Cytokinesis in Higher Plants
The Elaborate M Phase of Higher Organisms Evolved Gradually from Procaryotic Fission Mechanisms
Summary
References

Part V Cells in Their Social Context

CHAPTER 19 CELL JUNCTIONS, CELL ADHESION, AND THE EXTRACELLULAR MATRIX 1065

CELL JUNCTIONS

Occluding Junctions Form a Selective Permeability Barrier Across Epithelial Cell Sheets
Anchoring Junctions Connect the Cytoskeleton of a Cell Either to the Cytoskeleton of Its Neighbors or to the Extracellular Matrix
Adherens Junctions Connect Bundles of Actin Filaments from Cell to Cell
Desmosomes Connect Intermediate Filaments from Cell to Cell
Anchoring Junctions Formed by Integrins Bind Cells to the Extracellular Matrix: Focal Adhesions and Hemidesmosomes
Gap Junctions Allow Small Molecules to Pass Directly from Cell to Cell
A Gap-Junction Connexon Is Made Up of Six Transmembrane Connexin Subunits
Gap Junctions Have Diverse Functions
The Permeability of Gap Junctions Can Be Regulated
In Plants, Plasmodesmata Perform Many of the Same Functions as Gap Junctions
Summary

CELL–CELL ADHESION

Animal Cells Can Assemble into Tissues Either in Place or After They Migrate
Dissociated Vertebrate Cells Can Reassemble into Organized Tissues Through Selective Cell–Cell Adhesion
Cadherins Mediate Ca^{2+}-Dependent Cell–Cell Adhesion
Cadherins Have Crucial Roles in Development
Cadherins Mediate Cell–Cell Adhesion by a Homophilic Mechanism
Cadherins Are Linked to the Actin Cytoskeleton by Catenins

Selectins Mediate Transient Cell–Cell Adhesions in the Bloodstream
Members of the Immunoglobulin Superfamily of Proteins Mediate Ca^{2+}-Independent Cell–Cell Adhesion
Multiple Types of Cell-Surface Molecules Act in Parallel to Mediate Selective Cell–Cell Adhesion
Nonjunctional Contacts May Initiate Cell–Cell Adhesions That Junctional Contacts Then Orient and Stabilize
Summary

THE EXTRACELLULAR MATRIX OF ANIMALS

The Extracellular Matrix Is Made and Oriented by the Cells Within It
Glycosaminoglycan (GAG) Chains Occupy Large Amounts of Space and Form Hydrated Gels
Hyaluronan Is Thought to Facilitate Cell Migration During Tissue Morphogenesis and Repair
Proteoglycans Are Composed of GAG Chains Covalently Linked to a Core Protein
Proteoglycans Can Regulate the Activities of Secreted Proteins
GAG Chains May Be Highly Organized in the Extracellular Matrix
Cell-Surface Proteoglycans Act as Co-Receptors
Collagens Are the Major Proteins of the Extracellular Matrix
Collagens Are Secreted with a Nonhelical Extension at Each End
After Secretion, Fibrillar Procollagen Molecules Are Cleaved to Collagen Molecules, Which Assemble into Fibrils
Fibril-Associated Collagens Help Organize the Fibrils
Cells Help Organize the Collagen Fibrils They Secrete by Exerting Tension on the Matrix
Elastin Gives Tissues Their Elasticity
Fibronectin Is an Extracellular Protein That Helps Cells Attach to the Matrix
Fibronectin Exists in Both Soluble and Fibrillar Forms
Intracellular Actin Filaments Regulate the Assembly of Extracellular Fibronectin Fibrils
Glycoproteins in the Matrix Help Guide Cell Migration
Basal Laminae Are Composed Mainly of Type IV Collagen, Laminin, Nidogen, and a Heparan Sulfate Proteoglycan
Basal Laminae Perform Diverse Functions
The Extracellular Matrix Can Influence Cell Shape, Cell Survival, and Cell Proliferation
The Controlled Degradation of Matrix Components Helps Cells Migrate
Summary

INTEGRINS

Integrins Are Transmembrane Heterodimers
Integrins Must Interact with the Cytoskeleton to Bind Cells to the Extracellular Matrix
Cells Can Regulate the Activity of Their Integrins
Integrins Activate Intracellular Signaling Pathways
Summary

THE PLANT CELL WALL

The Composition of the Cell Wall Depends on the Cell Type
The Tensile Strength of the Cell Wall Allows Plant Cells to Develop Turgor Pressure
The Primary Cell Wall Is Built from Cellulose Microfibrils Interwoven with a Network of Pectic Polysaccharides
Microtubules Orient Cell-Wall Deposition
Summary
References

CHAPTER 20 GERM CELLS AND FERTILIZATION 1127

THE BENEFITS OF SEX

In Multicellular Animals and Most Plants the Diploid Phase Is Complex and Long, the Haploid Simple and Fleeting
Sexual Reproduction Gives a Competitive Advantage to Organisms in an Unpredictably Variable Environment
Summary

MEIOSIS

Duplicated Homologous Chromosomes Pair During Meiosis
Gametes Are Produced by Two Meiotic Cell Divisions
Genetic Reassortment Is Enhanced by Crossing-Over Between Homologous Nonsister Chromatids
Chiasmata Have an Important Role in Chromosome Segregation in Meiosis
Pairing of the Sex Chromosomes Ensures That They Also Segregate
Meiotic Chromosome Pairing Culminates in the Formation of the Synaptonemal Complex
Recombination Nodules Mark the Sites of Genetic Recombination
Genetic Maps Reveal Favored Sites for Crossovers
Meiosis Ends with Two Successive Cell Divisions Without DNA Replication
Summary

PRIMORDIAL GERM CELLS AND SEX DETERMINATION IN MAMMALS

Primordial Germ Cells Migrate into the Developing Gonad
The *Sry* Gene on the Y Chromosome Can Redirect a Female Embryo to Become a Male
Summary

EGGS

An Egg Is Highly Specialized for Independent Development, with Large Nutrient Reserves and an Elaborate Coat
Eggs Develop in Stages
Oocytes Use Special Mechanisms to Grow to Their Large Size
Summary

SPERM

Sperm Are Highly Adapted for Delivering Their DNA to an Egg
Sperm Are Produced Continuously in Most Mammals
Summary

FERTILIZATION

Species-Specific Binding to the Zona Pellucida Induces the Sperm to Undergo an Acrosome Reaction
The Egg Cortical Reaction Helps to Ensure That Only One Sperm Fertilizes the Egg
The Mechanism of Sperm–Egg Fusion Is Still Unknown
The Sperm Provides a Centriole for the Zygote
Summary
References

CHAPTER 21 DEVELOPMENT OF MULTICELLULAR ORGANISMS 1157

UNIVERSAL MECHANISMS OF ANIMAL DEVELOPMENT

Animals Share Some Basic Anatomical Features
Multicellular Animals are Enriched in Proteins Mediating Cell Interactions and Gene Regulation
Regulatory DNA Defines the Program of Development
Manipulation of the Embryo Reveals the Interactions Between its Cells
Studies of Mutant Animals Identify the Genes That Control Developmental Processes
A Cell Makes Developmental Decisions Long Before It Shows a Visible Change

Cells Have Remembered Positional Values That Reflect Their Location in the Body
Sister Cells Can Be Born Different by an Asymmetric Cell Division
Inductive Interactions Can Create Orderly Differences Between Initially Identical Cells.
Morphogens Are Long-Range Inducers that Exert Graded Effects
Extracellular Inhibitors of Signal Molecules Shape the Response to the Inducer
Programs That Are Intrinsic to a Cell Often Define the Time-Course of its Development
Initial Patterns are Established in Small Fields of Cells and Refined by Sequential Induction as the Embryo Grows
Summary

CAENORHABDITIS ELEGANS: DEVELOPMENT FROM THE PERSPECTIVE OF THE INDIVIDUAL CELL

Caenorhabditis elegans Is Anatomically Simple
Cell Fates in the Developing Nematode Are Almost Perfectly Predictable
Products of Maternal-Effect Genes Organize the Asymmetric Division of the Egg
Progressively More Complex Patterns Are Created by Cell–Cell Interactions
Microsurgery and Genetics Reveal the Logic of Developmental Control; Gene Cloning and Sequencing Reveal its Molecular Mechanisms
Cells Change Over Time in Their Responsiveness to Developmental Signals
Heterochronic Genes Control the Timing of Development
Cells Do Not Count Cell Divisions in Timing Their Internal Programs
Selected Cells Die by Apoptosis as Part of the Program of Development
Summary

DROSOPHILA AND THE MOLECULAR GENETICS OF PATTERN FORMATION: GENESIS OF THE BODY PLAN

The Insect Body is Constructed as a Series of Segmental Units
Drosophila Begins its Development as a Syncytium
Genetic Screens Define Groups of Genes Required for Specific Aspects of Early Patterning
Interactions of the Oocyte with its Surroundings Define the Axes of the Embryo: the Role of the Egg-Polarity Genes
The Dorsoventral Signaling Genes Create a Gradient of a Nuclear Gene Regulatory Protein
Dpp and Sog Set Up a Secondary Morphogen Gradient to Refine the Pattern of the Dorsal Part of the Embryo
The Insect Dorsoventral Axis Corresponds to the Vertebrate Ventrodorsal Axis
Three Classes of Segmentation Genes Refine the Anterior-Posterior Maternal Pattern and Subdivide the Embryo
The Localized Expression of Segmentation Genes Is Regulated by a Hierarchy of Positional Signals
The Modular Nature of Regulatory DNA Allows Genes to Have Multiple Independently Controlled Functions
Egg-Polarity, Gap, and Pair-Rule Genes Create a Transient Pattern That Is Remembered by Other Genes
Summary

HOMEOTIC SELECTOR GENES AND THE PATTERNING OF THE ANTEROPOSTERIOR AXIS

The HOX Code Specifies Anterior–Posterior Differences
Homeotic Selector Genes Code for DNA-Binding Proteins that Interact with Other Gene Regulatory Proteins
The Homeotic Selector Genes are Expressed Sequentially According to Their Order in the Hox Complex
The Hox Complex Carries a Permanent Record of Positional Information
The Anteroposterior Axis Is Controlled by Hox Selector Genes in Vertebrates Also
Summary

ORGANOGENESIS AND THE PATTERNING OF APPENDAGES

Conditional and Induced Somatic Mutations Make it Possible to Analyze Gene Functions Late in Development
Body Parts of the Adult Fly Develop from Imaginal Discs
Homeotic Selector Genes Are Essential for the Memory of Positional Information in Imaginal Disc Cells
Specific Regulatory Genes Define the Cells That Will Form an Appendage
The Insect Wing Disc is Divided into Compartments
Four Familiar Signaling Pathways Combine to Pattern the Wing Disc: Wingless, Hedgehog, Dpp and Notch
The Size of Each Compartment is Regulated by Interactions Among its Cells
Similar Mechanisms Pattern the Limbs of Vertebrates
Localized Expression of Specific Classes of Gene Regulatory Proteins Foreshadows Cell Differentiation
Lateral Inhibition Singles Out Sensory Mother Cells Within Proneural Clusters
Lateral Inhibition Drives the Progeny of the Sensory Mother Cell Toward Different Final Fates
Planar Polarity of Asymmetric Divisions is Controlled by Signaling via the Receptor Frizzled
Lateral Inhibition and Asymmetric Division Combine to Regulate Genesis of Neurons Throughout the Body
Notch Signaling Regulates the Fine-Grained Pattern of Differentiated Cell Types in Many Different Tissues
Some Key Regulatory Genes Define a Cell Type; Others Can Activate the Program for Creation of an Entire Organ
Summary

CELL MOVEMENTS AND THE SHAPING OF THE VERTEBRATE BODY

The Polarity of the Amphibian Embryo Depends on the Polarity of the Egg
Cleavage Produces Many Cells from One
Gastrulation Transforms a Hollow Ball of Cells into a Three-Layered Structure with a Primitive Gut
The Movements of Gastrulation Are Precisely Predictable
Chemical Signals Trigger the Mechanical Processes
Active Changes of Cell Packing Provide a Driving Force for Gastrulation
Changing Patterns of Cell Adhesion Molecules Force Cells Into New Arrangements
The Notochord Elongates, While the Neural Plate Rolls up to Form the Neural Tube
A Gene-Expression Oscillator Controls Segmentation of the Mesoderm Into Somites
Embryonic Tissues Are Invaded in a Strictly Controlled Fashion by Migratory Cells
The Distribution of Migrant Cells Depends on Survival Factors as Well as Guidance Cues
Left–Right Asymmetry of the Vertebrate Body Derives From Molecular Asymmetry in the Early Embryo
Summary

THE MOUSE

Mammalian Development Begins With a Specialized Preamble
The Early Mammalian Embryo Is Highly Regulative
Totipotent Embryonic Stem Cells Can Be Obtained from a Mammalian Embryo
Interactions Between Epithelium and Mesenchyme Generate Branching Tubular Structures
Summary

NEURAL DEVELOPMENT

Neurons Are Assigned Different Characters According to the Time and Place Where They Are Born
The Character Assigned to a Neuron at Its Birth Governs the Connections It Will Form
Each Axon or Dendrite Extends by Means of a Growth Cone at Its Tip
The Growth Cone Pilots the Developing Neurite Along a Precisely Defined Path *in vivo*

Growth Cones Can Change Their Sensibilities as They Travel
Target Tissues Release Neurotrophic Factors That Control Nerve Cell Growth and Survival
Neuronal Specificity Guides the Formation of Orderly Neural Maps
Axons from Different Regions of the Retina Respond Differently to a Gradient of Repulsive Molecules in the Tectum
Diffuse Patterns of Synaptic Connections Are Sharpened by Activity-Dependent Remodeling
Experience Molds the Pattern of Synaptic Connections in the Brain
Adult Memory and Developmental Synapse Remodeling May Depend on Similar Mechanisms
Summary

PLANT DEVELOPMENT

Arabidopsis Serves as a Model Organism for Plant Molecular Genetics
The *Arabidopsis* Genome Is Rich in Developmental Control Genes
Embryonic Development Starts by Establishing a Root–Shoot Axis and Then Halts Inside the Seed
The Parts of a Plant Are Generated Sequentially by Meristems
Development of the Seedling Depends on Environmental Signals
The Shaping of Each New Structure Depends on Oriented Cell Division and Expansion
Each Plant Module Grows From a Microscopic Set of Primordia in a Meristem
Cell Signaling Maintains the Meristem
Regulatory Mutations can Transform Plant Topology by Altering Cell Behavior in the Meristem
Long-Range Hormonal Signals Coordinate Developmental Events in Separate Parts of the Plant
Homeotic Selector Genes Specify the Parts of a Flower
Summary
References

CHAPTER 22 HISTOLOGY: THE LIVES AND DEATHS OF CELLS IN TISSUES 1259

EPIDERMIS AND ITS RENEWAL BY STEM CELLS

Epidermal Cells Form a Multilayered Waterproof Barrier
Differentiating Epidermal Cells Synthesize a Sequence of Different Keratins as They Mature
Epidermis Is Renewed by Stem Cells Lying in its Basal Layer
The Two Daughters of a Stem Cell Do Not Always Have to Become Different
The Basal Layer Contains Both Stem Cells and Transit Amplifying Cells
Epidermal Renewal Is Governed by Many Interacting Signals
The Mammary Gland Undergoes Cycles of Development and Regression
Summary

SENSORY EPITHELIA

Olfactory Sensory Neurons Are Continually Replaced
Auditory Hair Cells Have to Last a Lifetime
Most Permanent Cells Renew Their Parts: The Photoreceptor Cells of the Retina
Summary

THE AIRWAYS AND THE GUT

Adjacent Cell Types Collaborate in the Alveoli of the Lungs
Goblet Cells, Ciliated Cells, and Macrophages Collaborate to Keep the Airways Clean
The Lining of the Small Intestine Renews Itself Faster Than Any Other Tissue
Components of the Wnt Signaling Pathway Are Required to Maintain the Gut Stem-Cell Population
The Liver Functions as an Interface Between the Digestive Tract and the Blood
Liver Cell Loss Stimulates Liver Cell Proliferation
Summary

BLOOD VESSELS AND ENDOTHELIAL CELLS

Endothelial Cells Line All Blood Vessels
New Endothelial Cells Are Generated by Simple Duplication of Existing Endothelial Cells
New Capillaries Form by Sprouting
Angiogenesis is Controlled by Factors Released by the Surrounding Tissues
Summary

RENEWAL BY PLURIPOTENT STEM CELLS: BLOOD CELL FORMATION

The Three Main Categories of White Blood Cells: Granulocytes, Monocytes, and Lymphocytes
The Production of Each Type of Blood Cell in the Bone Marrow Is Individually Controlled
Bone Marrow Contains Hemopoietic Stem Cells
A Pluripotent Stem Cell Gives Rise to All Classes of Blood Cells
Commitment Is a Stepwise Process
The Number of Specialized Blood Cells Is Amplified by Divisions of Committed Progenitor Cells
Stem Cells Depend on Contact Signals From Stromal Cells
Factors That Regulate Hemopoiesis Can Be Analyzed in Culture
Erythropoiesis Depends on the Hormone Erythropoietin
Multiple CSFs Influence the Production of Neutrophils and Macrophages
The Behavior of a Hemopoietic Cell Depends Partly on Chance
Regulation of Cell Survival Is as Important as Regulation of Cell Proliferation
Summary

GENESIS, MODULATION, AND REGENERATION OF SKELETAL MUSCLE

New Skeletal Muscle Fibers Form by the Fusion of Myoblasts
Muscle Cells Can Vary Their Properties by Changing the Protein Isoforms They Contain
Skeletal Muscle Fibers Secrete Myostatin to Limit Their own Growth
Some Myoblasts Persist as Quiescent Stem Cells in the Adult
Summary

FIBROBLASTS AND THEIR TRANSFORMATIONS: THE CONNECTIVE-TISSUE CELL FAMILY

Fibroblasts Change Their Character in Response to Chemical Signals
The Extracellular Matrix May Influence Connective-Tissue Cell Differentiation by Affecting Cell Shape and Attachment
Fat Cells Can Develop from Fibroblasts
Leptin Secreted by Fat Cells Provides Negative Feedback to Inhibit Eating
Bone Is Continually Remodeled by the Cells Within It
Osteoblasts Secrete Bone Matrix, While Osteoclasts Erode It
During Development, Cartilage is Eroded by Osteoclasts to Make Way for Bone
Summary

STEM-CELL ENGINEERING

ES Cells Can Be Used to Make Any Part of the Body
Epidermal Stem Cell Populations Can Be Expanded in Culture for Tissue Repair
Neural Stem Cells Can Repopulate the Central Nervous System
The Stem Cells of Adult Tissues May Be More Versatile Than They Seem
Summary
References

Case No.: IPR2023-00322
U.S. Patent No. 10,273,492

CHAPTER 23 CANCER 1313

CANCER AS A MICROEVOLUTIONARY PROCESS

Cancer Cells Reproduce Without Restraint and Colonize Foreign Tissues
Most Cancers Derive From a Single Abnormal Cell
Cancers Result From Somatic Mutation
A Single Mutation Is Not Enough to Cause Cancer
Cancers Develop in Slow Stages From Mildly Aberrant Cells
Tumor Progression Involves Successive Rounds of Mutation and Natural Selection
Many Human Cancer Cells Are Genetically Unstable
Cancerous Growth Often Depends on Defective Control of Cell Death or Cell Differentiation
Many Cancer Cells Escape a Built-in Limit to Cell Proliferation
To Metastasize, Malignant Cancer Cells Must Survive and Proliferate in an Alien Environment
Six Key Properties Make Cells Capable of Cancerous Growth
Summary

THE PREVENTABLE CAUSES OF CANCER

Many, But Not All, Cancer-Causing Agents Damage DNA
The Development of a Cancer Can Be Promoted by Factors That Do Not Alter the Cell's DNA Sequence
Viruses and Other Infections Contribute to a Significant Proportion of Human Cancers
Identification of Carcinogens Reveals Ways to Avoid Cancer
Summary

FINDING THE CANCER-CRITICAL GENES

Different Methods Are Used to Identify Gain-of-Function and Loss-of-Function Mutations
Oncogenes Are Identified Through Their Dominant Transforming Effects
Tumor Suppressor Genes Can Sometimes Be Identified by Study of Rare Hereditary Cancer Syndromes
Tumor Suppressor Genes Can Be Identified Even Without Clues from Heritable Cancer Syndromes
Genes Mutated in Cancer Can Be Made Overactive or Underactive in Many Ways
The Hunt for Cancer-Critical Genes Continues
Summary

THE MOLECULAR BASIS OF CANCER-CELL BEHAVIOR

Studies of Developing Embryos and Transgenic Mice Help to Uncover the Function of Cancer-Critical Genes
Many Cancer-Critical Genes Regulate Cell Division
Mutations in Genes that Regulate Apoptosis Allow Cancer Cells to Escape Suicide
Mutations in the *p53* Gene Allow Cancer Cells to Survive and Proliferate Despite DNA Damage
DNA Tumor Viruses Activate the Cell's Replication Machinery by Blocking the Action of Key Tumor Suppressor Genes
Telomere Shortening May Pave the Way to Cancer in Humans
In a Population of Telomere-Deficient Cells, Loss of p53 Opens an Easy Gateway to Cancer
The Mutations that Lead to Metastasis Are Still a Mystery
Colorectal Cancers Evolve Slowly Via a Succession of Visible Changes
A Few Key Genetic Lesions Are Common to a Majority of Cases of Colorectal Cancer
Defects in DNA Mismatch Repair Provide an Alternative Route to Colorectal Cancer
The Steps of Tumor Progression Can Be Correlated with Specific Mutations
Each Case of Cancer Is Characterized by Its Own Array of Genetic Lesions
Summary

CANCER TREATMENT: PRESENT AND FUTURE

The Search for Cancer Cures Is Difficult but Not Hopeless
Current Therapies Exploit the Loss of Cell-Cycle Control and the Genetic Instability of Cancer Cells
Cancers Can Evolve Resistance to Therapies
New Therapies May Emerge from Our Knowledge of Cancer Biology
Treatments Can Be Designed to Attack Cells that Lack p53
Tumor Growth Can Be Choked by Depriving the Cancer Cells of Their Blood Supply
Small Molecules Can Be Designed to Target Specific Oncogenic Proteins
Understanding of Cancer Biology Leads Toward Rational, Tailored Medical Treatments
Summary
References

CHAPTER 24 THE ADAPTIVE IMMUNE SYSTEM 1363

LYMPHOCYTES AND THE CELLULAR BASIS OF ADAPTIVE IMMUNITY

Lymphocytes Are Required for Adaptive Immunity
The Innate and Adaptive Immune Systems Work Together
B Lymphocytes Develop in the Bone Marrow; T Lymphocytes Develop in the Thymus
The Adaptive Immune System Works by Clonal Selection
Most Antigens Activate Many Different Lymphocyte Clones
Immunological Memory Is Due to Both Clonal Expansion and Lymphocyte Differentiation
Acquired Immunological Tolerance Ensures That Self Antigens Are Not Attacked
Lymphocytes Continuously Circulate Through Peripheral Lymphoid Organs
Summary

B CELLS AND ANTIBODIES

B Cells Make Antibodies as Both Cell-Surface Receptors and Secreted Molecules
A Typical Antibody Has Two Identical Antigen-Binding Sites
An Antibody Molecule Is Composed of Heavy and Light Chains
There Are Five Classes of Heavy Chains, Each With Different Biological Properties
The Strength of an Antibody–Antigen Interaction Depends on Both the Number and the Affinity of the Antigen-Binding Sites
Light and Heavy Chains Consist of Constant and Variable Regions
The Light and Heavy Chains Are Composed of Repeating Ig Domains
An Antigen-Binding Site Is Constructed from Hypervariable Loops
Summary

THE GENERATION OF ANTIBODY DIVERSITY

Antibody Genes Are Assembled from Separate Gene Segments During B Cell Development
Each Variable Region Is Encoded by More Than One Gene Segment
Imprecise Joining of Gene Segments Greatly Increases the Diversity of V Regions
Antigen-driven Somatic Hypermutation Fine-Tunes Antibody Responses
The Control of V(D)J Joining Ensures That B Cells Are Monospecific
When Activated by Antigen, a B Cell Switches from Making a Membrane-Bound Antibody to Making a Secreted Form of the Same Antibody
B Cells Can Switch the Class of Antibody They Make
Summary

Motif Exhibit 1059, 495 of 653

T CELLS AND MHC PROTEINS

T Cell Receptors Are Antibodylike Heterodimers
Antigen-Presenting Cells Activate T Cells
Effector Cytotoxic T Cells Induce Infected Target Cells to Kill Themselves
Effector Helper T Cells Help Activate Macrophages, B Cells, and Cytotoxic T Cells
T Cells Recognize Foreign Peptides Bound to MHC Proteins
MHC Proteins Were Identified in Transplantation Reactions Before Their Functions Were Known
Class I and Class II MHC Proteins Are Structurally Similar Heterodimers
An MHC Protein Binds a Peptide and Interacts with a T Cell Receptor
MHC Proteins Help Direct T Cells to Their Appropriate Targets
CD4 and CD8 Co-receptors Bind to Nonvariable Parts of MHC Proteins
Cytotoxic T Cells Recognize Fragments of Foreign Cytosolic Proteins in Association with Class I MHC Proteins
Helper T Cells Recognize Fragments of Endocytosed Foreign Protein Associated with Class II MHC Proteins
Potentially Useful T Cells Are Positively Selected in the Thymus
Many Developing T Cells That Could Be Activated By Self Peptides Are Eliminated in the Thymus
The Function of MHC Proteins Explains Their Polymorphism
Summary

HELPER T CELLS AND LYMPHOCYTE ACTIVATION

Costimulatory Proteins on Antigen-Presenting Cells Help Activate T Cells
The Subclass of Helper T Cell Determines the Nature of the Adaptive Immune Response
T_H1 Cells Help Activate Macrophages at Sites of Infection
Antigen Binding Provides Signal 1 to B Cells
Helper T Cells Provide Signal 2 to B Cells
Immune Recognition Molecules Belong to an Ancient Superfamily
Summary
References

CHAPTER 25 PATHOGENS, INFECTION, AND INNATE IMMUNITY 1423

INTRODUCTION TO PATHOGENS

Pathogens Have Evolved Specific Mechanisms for Interacting with Their Hosts
The Signs and Symptoms of Infection May Be Caused Either by the Pathogen or by the Host's Responses
Pathogens Are Phylogenetically Diverse
Bacterial Pathogens Carry Specialized Virulence Genes
Fungal and Protozoan Parasites Have Complex Life Cycles with Multiple Forms
Viruses Exploit Host Cell Machinery for All Aspects of Their Multiplication
Prions Are Infectious Proteins
Summary

CELL BIOLOGY OF INFECTION

Pathogens Cross Protective Barriers to Colonize the Host
Pathogens That Colonize Epithelia Must Avoid Clearance by the Host
Intracellular Pathogens Have Mechanisms for Both Entering and Leaving Host Cells
Viruses Bind to Molecules Displayed on the Host Cell Surface
Viruses Enter Host Cells By Membrane Fusion, Pore Formation, or Membrane Disruption
Bacteria Enter Host Cells by Phagocytosis
Intracellular Parasites Actively Invade Host Cells
Many Pathogens Alter Membrane Traffic in the Host Cell
Viruses and Bacteria Exploit the Host Cell Cytoskeleton for Intracellular Movement
Viruses Take Over the Metabolism of the Host Cell
Pathogens Can Alter the Behavior of the Host Organism to Facilitate the Spread of the Pathogen
Pathogens Evolve Rapidly
Drug Resistant Pathogens Are a Growing Problem
Summary

INNATE IMMUNITY

Epithelial Surfaces Help Prevent Infection
Human Cells Recognize Conserved Features of Pathogens
Complement Activation Targets Pathogens for Phagocytosis or Lysis
Toll-like Proteins Are an Ancient Family of Pattern Recognition Receptors
Phagocytic Cells Seek, Engulf, and Destroy Pathogens
Activated Macrophages Recruit Additional Phagocytic Cells to Sites of Infection
Virus-Infected Cells Take Drastic Measures To Prevent Viral Replication
Natural Killer Cells Induce Virus-Infected Cells to Kill Themselves
Summary
References

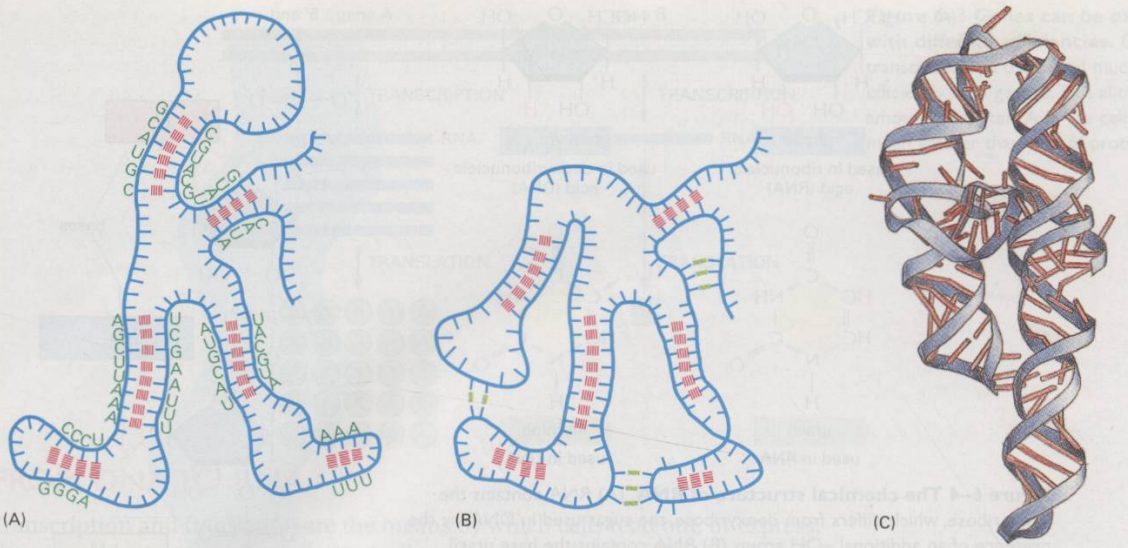


Figure 6–6 RNA can fold into specific structures. RNA is largely single-stranded, but it often contains short stretches of nucleotides that can form conventional base-pairs with complementary sequences found elsewhere on the same molecule. These interactions, along with additional "nonconventional" base-pair interactions, allow an RNA molecule to fold into a three-dimensional structure that is determined by its sequence of nucleotides. (A) Diagram of a folded RNA structure showing only conventional base-pair interactions; (B) structure with both conventional *(red)* and nonconventional *(green)* base-pair interactions; (C) structure of an actual RNA, a portion of a group I intron (see Figure 6–36). Each conventional base-pair interaction is indicated by a "rung" in the double helix. Bases in other configurations are indicated by broken rungs.

relatively short time, the synthesis of additional RNA molecules being started before the first RNA is completed (Figure 6–9). When RNA polymerase molecules follow hard on each other's heels in this way, each moving at about 20 nucleotides per second (the speed in eucaryotes), over a thousand transcripts can be synthesized in an hour from a single gene.

Although RNA polymerase catalyzes essentially the same chemical reaction as DNA polymerase, there are some important differences between the two enzymes. First, and most obvious, RNA polymerase catalyzes the linkage of ribonucleotides, not deoxyribonucleotides. Second, unlike the DNA polymerases involved in DNA replication, RNA polymerases can start an RNA chain without a primer. This difference may exist because transcription need not be as accurate as DNA replication (see Table 5–1, p. 243). Unlike DNA, RNA does not permanently store genetic information in cells. RNA polymerases make about one mistake for every 10^4 nucleotides copied into RNA (compared with an error rate for direct copying by DNA polymerase of about one in 10^7 nucleotides), and the consequences of an error in RNA transcription are much less significant than that in DNA replication.

Although RNA polymerases are not nearly as accurate as the DNA polymerases that replicate DNA, they nonetheless have a modest proofreading mechanism. If the incorrect ribonucleotide is added to the growing RNA chain, the polymerase can back up, and the active site of the enzyme can perform an excision reaction that mimics the reverse of the polymerization reaction, except that water instead of pyrophosphate is used (see Figure 5–4). RNA polymerase hovers around a misincorporated ribonucleotide longer than it does for a correct addition, causing excision to be favored for incorrect nucleotides. However, RNA polymerase also excises many correct bases as part of the cost for improved accuracy.

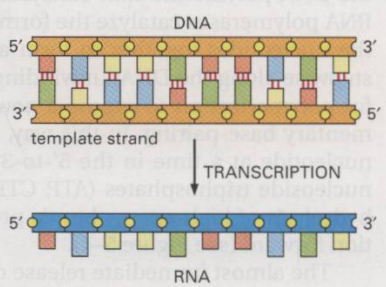


Figure 6–7 DNA transcription produces a single-stranded RNA molecule that is complementary to one strand of DNA.

Cells Produce Several Types of RNA

The majority of genes carried in a cell's DNA specify the amino acid sequence of proteins; the RNA molecules that are copied from these genes (which ultimately direct the synthesis of proteins) are called **messenger RNA (mRNA)** molecules.

This material may be protected by Copyright law (Title 17 U.S. Code)

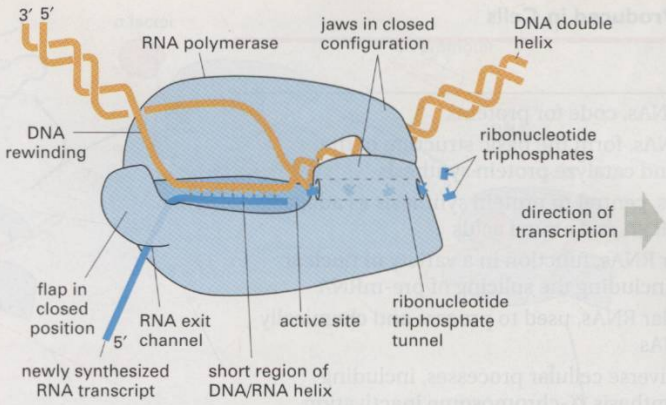


Figure 6–8 DNA is transcribed by the enzyme RNA polymerase. The RNA polymerase *(pale blue)* moves stepwise along the DNA, unwinding the DNA helix at its active site. As it progresses, the polymerase adds nucleotides (here, *small "T" shapes)* one by one to the RNA chain at the polymerization site using an exposed DNA strand as a template. The RNA transcript is thus a single-stranded complementary copy of one of the two DNA strands. The polymerase has a rudder (see Figure 6–11) that displaces the newly formed RNA, allowing the two strands of DNA behind the polymerase to rewind. A short region of DNA/RNA helix (approximately nine nucleotides in length) is therefore formed only transiently, and a "window" of DNA/RNA helix therefore moves along the DNA with the polymerase. The incoming nucleotides are in the form of ribonucleoside triphosphates (ATP, UTP, CTP, and GTP), and the energy stored in their phosphate–phosphate bonds provides the driving force for the polymerization reaction (see Figure 5–4). (Adapted from a figure kindly supplied by Robert Landick.)

The final product of a minority of genes, however, is the RNA itself. Careful analysis of the complete DNA sequence of the genome of the yeast *S. cerevisiae* has uncovered well over 750 genes (somewhat more than 10% of the total number of yeast genes) that produce RNA as their final product, although this number includes multiple copies of some highly repeated genes. These RNAs, like proteins, serve as enzymatic and structural components for a wide variety of processes in the cell. In Chapter 5 we encountered one of those RNAs, the template carried by the enzyme telomerase. Although not all of their functions are known, we see in this chapter that some *small nuclear RNA* (*snRNA*) molecules direct the splicing of pre-mRNA to form mRNA, that *ribosomal RNA (rRNA)* molecules form the core of ribosomes, and that *transfer RNA (tRNA)* molecules form the adaptors that select amino acids and hold them in place on a ribosome for incorporation into protein (Table 6–1).

Each transcribed segment of DNA is called a *transcription unit*. In eucaryotes, a transcription unit typically carries the information of just one gene, and therefore codes for either a single RNA molecule or a single protein (or group of related proteins if the initial RNA transcript is spliced in more than one way to produce different mRNAs). In bacteria, a set of adjacent genes is often transcribed as a unit; the resulting mRNA molecule therefore carries the information for several distinct proteins.

Overall, RNA makes up a few percent of a cell's dry weight. Most of the RNA in cells is rRNA; mRNA comprises only 3–5% of the total RNA in a typical mammalian cell. The mRNA population is made up of tens of thousands of different species, and there are on average only 10–15 molecules of each species of mRNA present in each cell.

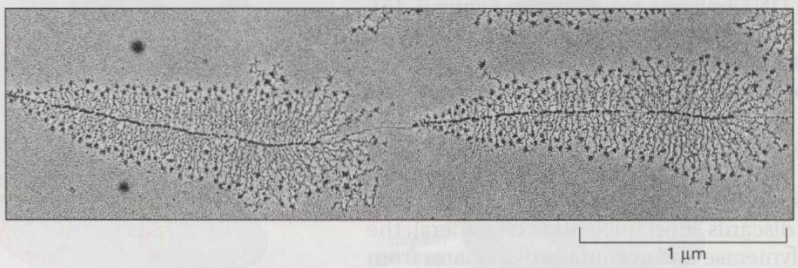


Figure 6–9 Transcription of two genes as observed under the electron microscope. The micrograph shows many molecules of RNA polymerase simultaneously transcribing each of two adjacent genes. Molecules of RNA polymerase are visible as a series of dots along the DNA with the newly synthesized transcripts (fine threads) attached to them. The RNA molecules (ribosomal RNAs) shown in this example are not translated into protein but are instead used directly as components of ribosomes, the machines on which translation takes place. The particles at the 5′ end (the free end) of each rRNA transcript are believed to reflect the beginnings of ribosome assembly. From the lengths of the newly synthesized transcripts, it can be deduced that the RNA polymerase molecules are transcribing from left to right. (Courtesy of Ulrich Scheer.)

TABLE 6–1 Principal Types of RNAs Produced in Cells

TYPE OF RNA	FUNCTION
mRNAs	messenger RNAs, code for proteins
rRNAs	ribosomal RNAs, form the basic structure of the ribosome and catalyze protein synthesis
tRNAs	transfer RNAs, central to protein synthesis as adaptors between mRNA and amino acids
snRNAs	small nuclear RNAs, function in a variety of nuclear processes, including the splicing of pre-mRNA
snoRNAs	small nucleolar RNAs, used to process and chemically modify rRNAs
Other noncoding RNAs	function in diverse cellular processes, including telomere synthesis, X-chromosome inactivation, and the transport of proteins into the ER

Signals Encoded in DNA Tell RNA Polymerase Where to Start and Stop

To transcribe a gene accurately, RNA polymerase must recognize where on the genome to start and where to finish. The way in which RNA polymerases perform these tasks differs somewhat between bacteria and eucaryotes. Because the process in bacteria is simpler, we look there first.

The initiation of transcription is an especially important step in gene expression because it is the main point at which the cell regulates which proteins are to be produced and at what rate. Bacterial RNA polymerase is a multisubunit complex. A detachable subunit, called *sigma (σ) factor*, is largely responsible for its ability to read the signals in the DNA that tell it where to begin transcribing (Figure 6–10). RNA polymerase molecules adhere only weakly to the bacterial DNA when they collide with it, and a polymerase molecule typically slides rapidly along the long DNA molecule until it dissociates again. However, when the polymerase slides into a region on the DNA double helix called a **promoter**, a special sequence of nucleotides indicating the starting point for RNA synthesis, it binds tightly to it. The polymerase, using its σ factor, recognizes this DNA sequence by making specific contacts with the portions of the bases that are exposed on the outside of the helix *(Step 1* in Figure 6–10).

After the RNA polymerase binds tightly to the promoter DNA in this way, it opens up the double helix to expose a short stretch of nucleotides on each strand *(Step 2* in Figure 6–10). Unlike a DNA helicase reaction (see Figure 5–15), this limited opening of the helix does not require the energy of ATP hydrolysis. Instead, the polymerase and DNA both undergo reversible structural changes that result in a more energetically favorable state. With the DNA unwound, one of the two exposed DNA strands acts as a template for complementary base-pairing with incoming ribonucleotides (see Figure 6–7), two of which are joined together by the polymerase to begin an RNA chain. After the first ten or so nucleotides of RNA have been synthesized (a relatively inefficient process during which polymerase synthesizes and discards short nucleotide oligomers), the σ factor relaxes its tight hold on the polymerase and evenutally dissociates from it. During this process, the polymerase undergoes additional structural changes that enable it to move forward rapidly, transcribing without the σ factor *(Step 4* in Figure 6–10). Chain elongation continues (at a speed of approximately 50 nucleotides/sec for bacterial RNA polymerases) until the enzyme encounters a second signal in the DNA, the **terminator** (described below), where the polymerase halts and releases both the DNA template and the newly made RNA chain *(Step 7* in Figure 6–10). After the polymerase has been released at a terminator, it reassociates with a free σ factor and searches for a new promoter, where it can begin the process of transcription again.

Several structural features of bacterial RNA polymerase make it particularly adept at performing the transcription cycle just described. Once the σ factor

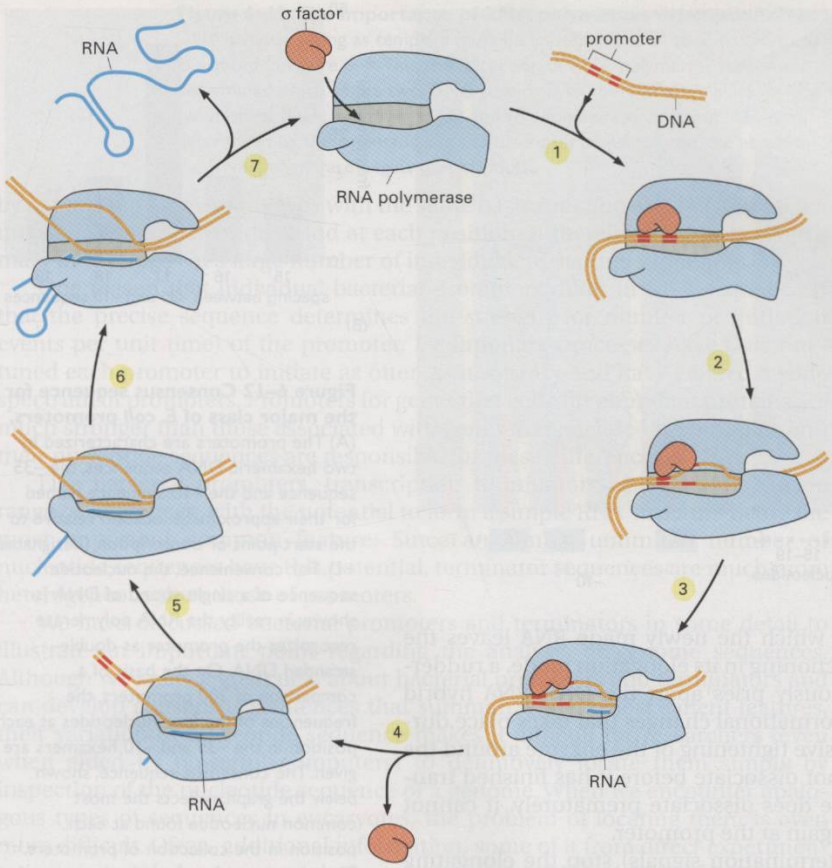


Figure 6–10 The transcription cycle of bacterial RNA polymerase. In step 1, the RNA polymerase holoenzyme (core polymerase plus σ factor) forms and then locates a promoter (see Figure 6–12). The polymerase unwinds the DNA at the position at which transcription is to begin (step 2) and begins transcribing (step 3). This initial RNA synthesis (sometimes called "abortive initiation") is relatively inefficient. However, once RNA polymerase has managed to synthesize about 10 nucleotides of RNA, σ relaxes its grip, and the polymerase undergoes a series of conformational changes (which probably includes a tightening of its jaws and the placement of RNA in the exit channel [see Figure 6–11]). The polymerase now shifts to the elongation mode of RNA synthesis (step 4), moving rightwards along the DNA in this diagram. During the elongation mode (step 5) transcription is highly processive, with the polymerase leaving the DNA template and releasing the newly transcribed RNA only when it encounters a termination signal (step 6). Termination signals are encoded in DNA and many function by forming an RNA structure that destabilizes the polymerase's hold on the RNA, as shown here. In bacteria, all RNA molecules are synthesized by a single type of RNA polymerase and the cycle depicted in the figure therefore applies to the production of mRNAs as well as structural and catalytic RNAs. (Adapted from a figure kindly supplied by Robert Landick.)

positions the polymerase on the promoter and the template DNA has been unwound and pushed to the active site, a pair of moveable jaws is thought to clamp onto the DNA (Figure 6–11). When the first 10 nucleotides have been transcribed, the dissociation of σ allows a flap at the back of the polymerase to

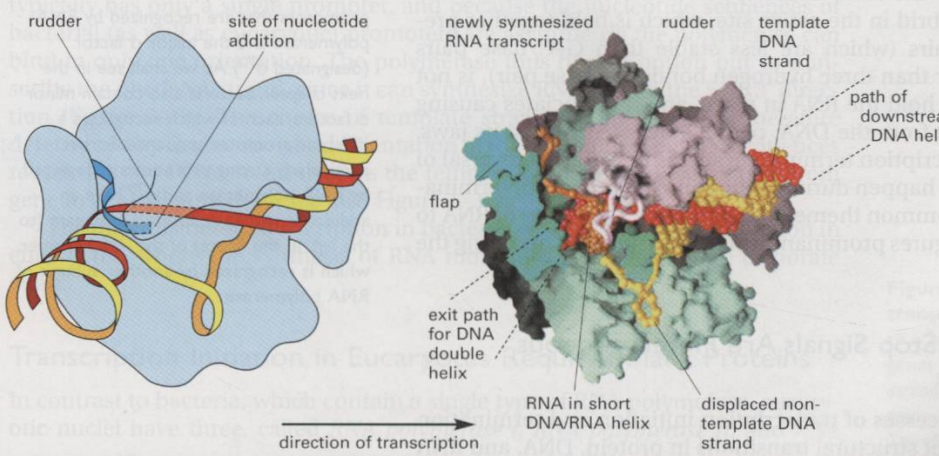


Figure 6–11 The structure of a bacterial RNA polymerase. Two depictions of the three-dimensional structure of a bacterial RNA polymerase, with the DNA and RNA modeled in. This RNA polymerase is formed from four different subunits, indicated by different colors *(right)*. The DNA strand used as a template is *red*, and the non-template strand is *yellow*. The rudder wedges apart the DNA–RNA hybrid as the polymerase moves. For simplicity only the polypeptide backbone of the rudder is shown in the right-hand figure, and the DNA exiting from the polymerase has been omitted. Because the RNA polymerase is depicted in the elongation mode, the σ factor is absent. (Courtesy of Seth Darst.)

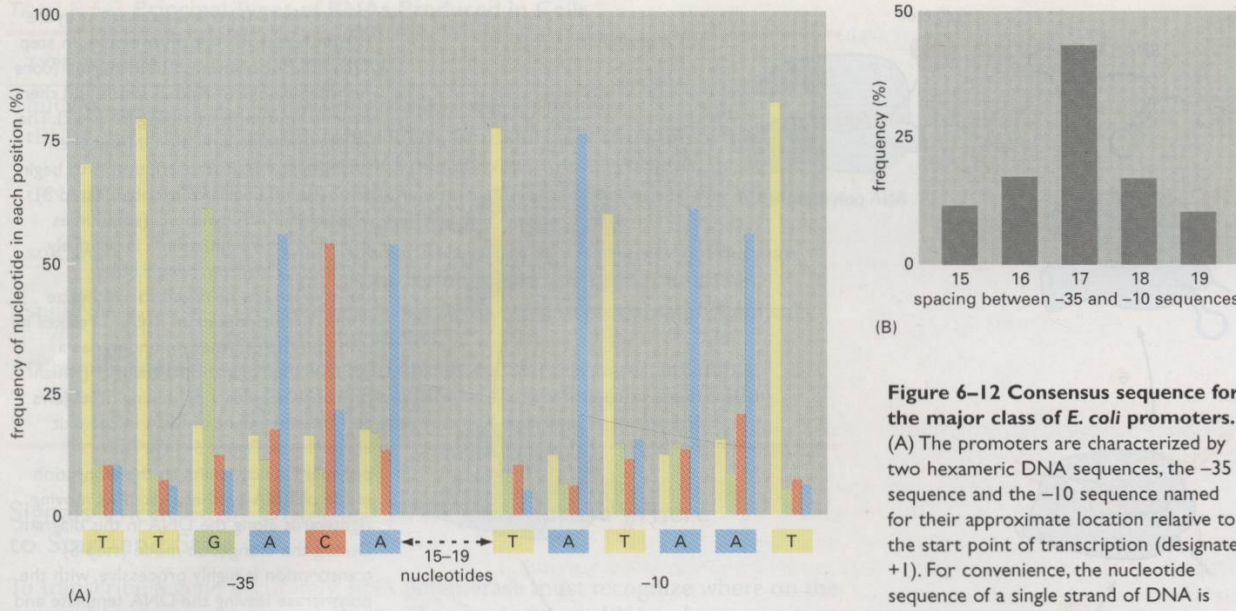


Figure 6–12 Consensus sequence for the major class of *E. coli* promoters. (A) The promoters are characterized by two hexameric DNA sequences, the –35 sequence and the –10 sequence named for their approximate location relative to the start point of transcription (designated +1). For convenience, the nucleotide sequence of a single strand of DNA is shown; in reality the RNA polymerase recognizes the promoter as double-stranded DNA. On the basis of a comparison of 300 promoters, the frequencies of the four nucleotides at each position in the –35 and –10 hexamers are given. The consensus sequence, shown *below* the graph, reflects the most common nucleotide found at each position in the collection of promoters. The sequence of nucleotides between the –35 and –10 hexamers shows no significant similarities among promoters. (B) The distribution of spacing between the –35 and –10 hexamers found in *E. coli* promoters. The information displayed in these two graphs applies to *E. coli* promoters that are recognized by RNA polymerase and the major σ factor (designated σ^{70}). As we shall see in the next chapter, bacteria also contain minor σ factors, each of which recognizes a different promoter sequence. Some particularly strong promoters recognized by RNA polymerase and σ^{70} have an additional sequence, located upstream (to the *left*, in the figure) of the –35 hexamer, which is recognized by another subunit of RNA polymerase.

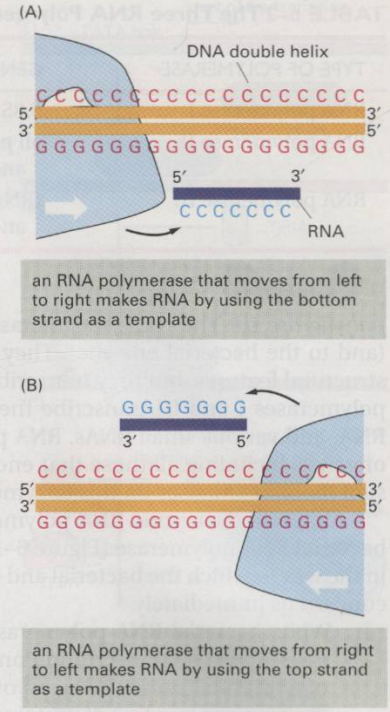
close to form an exit tunnel through which the newly made RNA leaves the enzyme. With the polymerase now functioning in its elongation mode, a rudder-like structure in the enzyme continuously pries apart the DNA–RNA hybrid formed. We can view the series of conformational changes that takes place during transcription initiation as a successive tightening of the enzyme around the DNA and RNA to ensure that it does not dissociate before it has finished transcribing a gene. If an RNA polymerase does dissociate prematurely, it cannot resume synthesis but must start over again at the promoter.

How do the signals in the DNA (termination signals) stop the elongating polymerase? For most bacterial genes a termination signal consists of a string of A–T nucleotide pairs preceded by a two-fold symmetric DNA sequence, which, when transcribed into RNA, folds into a "hairpin" structure through Watson–Crick base-pairing (see Figure 6–10). As the polymerase transcribes across a terminator, the hairpin may help to wedge open the movable flap on the RNA polymerase and release the RNA transcript from the exit tunnel. At the same time, the DNA–RNA hybrid in the active site, which is held together predominantly by U–A base pairs (which are less stable than G–C base pairs because they form two rather than three hydrogen bonds per base pair), is not sufficiently strong enough to hold the RNA in place, and it dissociates causing the release of the polymerase from the DNA, perhaps by forcing open its jaws. Thus, in some respects, transcription termination seems to involve a reversal of the structural transitions that happen during initiation. The process of termination also is an example of a common theme in this chapter: the ability of RNA to fold into specific structures figures prominantly in many aspects of decoding the genome.

Transcription Start and Stop Signals Are Heterogeneous in Nucleotide Sequence

As we have just seen, the processes of transcription initiation and termination involve a complicated series of structural transitions in protein, DNA, and RNA molecules. It is perhaps not surprising that the signals encoded in DNA that specify these transitions are difficult for researchers to recognize. Indeed, a comparison of many different bacterial promoters reveals that they are heterogeneous in DNA sequence. Nevertheless, they all contain related sequences, reflecting in part aspects of the DNA that are recognized directly by the σ factor. These common features are often summarized in the form of a *consensus sequence* (Figure 6–12). In general, a consensus nucleotide sequence is derived

Figure 6–13 The importance of RNA polymerase orientation. The DNA strand serving as template must be traversed in a 3′ to 5′ direction, as illustrated in Figure 6–9. Thus, the direction of RNA polymerase movement determines which of the two DNA strands is to serve as a template for the synthesis of RNA, as shown in (A) and (B). Polymerase direction is, in turn, determined by the orientation of the promoter sequence, the site at which the RNA polymerase begins transcription.



by comparing many sequences with the same basic function and tallying up the most common nucleotide found at each position. It therefore serves as a summary or "average" of a large number of individual nucleotide sequences.

One reason that individual bacterial promoters differ in DNA sequence is that the precise sequence determines the strength (or number of initiation events per unit time) of the promoter. Evolutionary processes have thus fine-tuned each promoter to initiate as often as necessary and have created a wide spectrum of promoters. Promoters for genes that code for abundant proteins are much stronger than those associated with genes that encode rare proteins, and their nucleotide sequences are responsible for these differences.

Like bacterial promoters, transcription terminators also include a wide range of sequences, with the potential to form a simple RNA structure being the most important common feature. Since an almost unlimited number of nucleotide sequences have this potential, terminator sequences are much more heterogeneous than those of promoters.

We have discussed bacterial promoters and terminators in some detail to illustrate an important point regarding the analysis of genome sequences. Although we know a great deal about bacterial promoters and terminators and can develop consensus sequences that summarize their most salient features, their variation in nucleotide sequence makes it difficult for researchers (even when aided by powerful computers) to definitively locate them simply by inspection of the nucleotide sequence of a genome. When we encounter analogous types of sequences in eucaryotes, the problem of locating them is even more difficult. Often, additional information, some of it from direct experimentation, is needed to accurately locate the short DNA signals contained in genomes.

Promoter sequences are asymmetric (see Figure 6–12), and this feature has important consequences for their arrangement in genomes. Since DNA is double-stranded, two different RNA molecules could in principle be transcribed from any gene, using each of the two DNA strands as a template. However a gene typically has only a single promoter, and because the nucleotide sequences of bacterial (as well as eucaryotic) promoters are asymmetric the polymerase can bind in only one orientation. The polymerase thus has no option but to transcribe the one DNA strand, since it can synthesize RNA only in the 5′ to 3′ direction (Figure 6–13). The choice of template strand for each gene is therefore determined by the location and orientation of the promoter. Genome sequences reveal that the DNA strand used as the template for RNA synthesis varies from gene to gene (Figure 6–14; see also Figure 1–31).

Having considered transcription in bacteria, we now turn to the situation in eucaryotes, where the synthesis of RNA molecules is a much more elaborate affair.

Transcription Initiation in Eucaryotes Requires Many Proteins

In contrast to bacteria, which contain a single type of RNA polymerase, eucaryotic nuclei have three, called *RNA polymerase I, RNA polymerase II,* and *RNA*

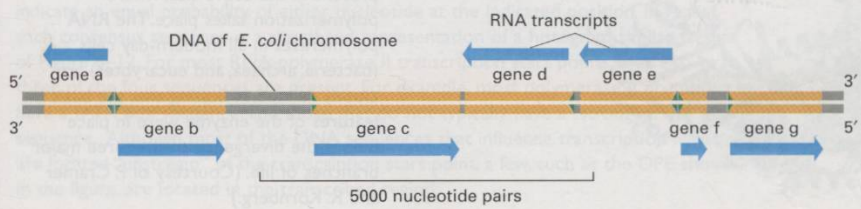


Figure 6–14 Directions of transcription along a short portion of a bacterial chromosome. Some genes are transcribed using one DNA strand as a template, while others are transcribed using the other DNA strand. The direction of transcription is determined by the promoter at the beginning of each gene *(green arrowheads).* Approximately 0.2% (9000 base pairs) of the *E. coli* chromosome is depicted here. The genes transcribed from *left* to *right* use the bottom DNA strand as the template; those transcribed from *right* to *left* use the top strand as the template.

TABLE 6–2 The Three RNA Polymerases in Eucaryotic Cells

TYPE OF POLYMERASE	GENES TRANSCRIBED
RNA polymerase I	5.8S, 18S, and 28S rRNA genes
RNA polymerase II	all protein-coding genes, plus snoRNA genes and some snRNA genes
RNA polymerase III	tRNA genes, 5S rRNA genes, some snRNA genes and genes for other small RNAs

polymerase III. The three polymerases are structurally similar to one another (and to the bacterial enzyme). They share some common subunits and many structural features, but they transcribe different types of genes (Table 6–2). RNA polymerases I and III transcribe the genes encoding transfer RNA, ribosomal RNA, and various small RNAs. RNA polymerase II transcribes the vast majority of genes, including all those that encode proteins, and our subsequent discussion therefore focuses on this enzyme.

Although eucaryotic RNA polymerase II has many structural similarities to bacterial RNA polymerase (Figure 6–15), there are several important differences in the way in which the bacterial and eucaryotic enzymes function, two of which concern us immediately.

1. While bacterial RNA polymerase (with σ factor as one of its subunits) is able to initiate transcription on a DNA template *in vitro* without the help of additional proteins, eucaryotic RNA polymerases cannot. They require the help of a large set of proteins called *general transcription factors*, which must assemble at the promoter with the polymerase before the polymerase can begin transcription.
2. Eucaryotic transcription initiation must deal with the packing of DNA into nucleosomes and higher order forms of chromatin structure, features absent from bacterial chromosomes.

RNA Polymerase II Requires General Transcription Factors

The discovery that, unlike bacterial RNA polymerase, purified eucaryotic RNA polymerase II could not initiate transcription *in vitro* led to the discovery and purification of the additional factors required for this process. These **general transcription factors** help to position the RNA polymerase correctly at the promoter, aid in pulling apart the two strands of DNA to allow transcription to begin, and release RNA polymerase from the promoter into the elongation mode once transcription has begun. The proteins are "general" because they assemble on all promoters used by RNA polymerase II; consisting of a set of interacting proteins, they are designated as *TFII* (for transcription factor for polymerase II),



Figure 6–15 Structural similarity between a bacterial RNA polymerase and a eucaryotic RNA polymerase II. Regions of the two RNA polymerases that have similar structures are indicated in *green*. The eucaryotic polymerase is larger than the bacterial enzyme (12 subunits instead of 5), and some of the additional regions are shown in *gray*. The *blue* spheres represent Zn atoms that serve as structural components of the polymerases, and the *red* sphere represents the Mg atom present at the active site, where polymerization takes place. The RNA polymerases in all modern-day cells (bacteria, archaea, and eucaryotes) are closely related, indicating that the basic features of the enzyme were in place before the divergence of the three major branches of life. (Courtesy of P. Cramer and R. Kornberg.)

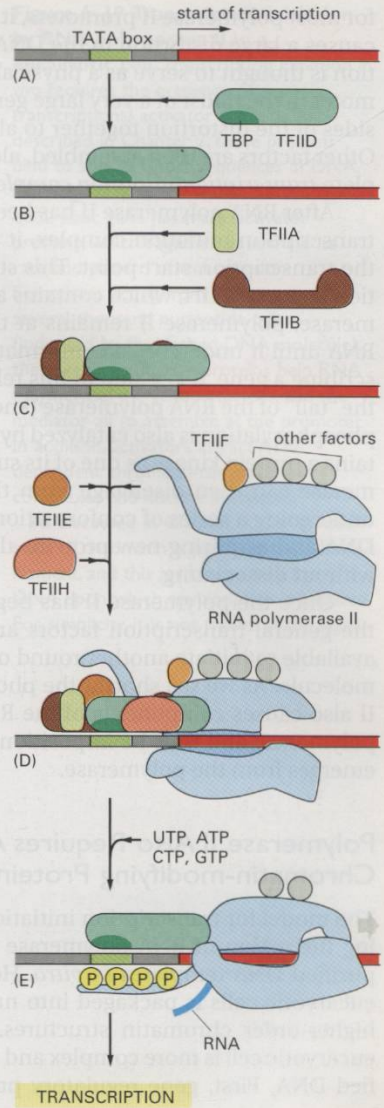
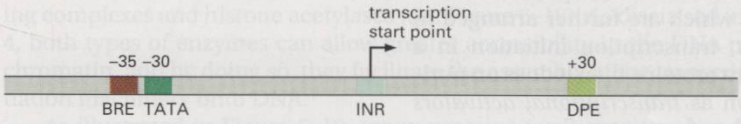


Figure 6–16 Initiation of transcription of a eucaryotic gene by RNA polymerase II. To begin transcription, RNA polymerase requires a number of general transcription factors (called TFIIA, TFIIB, and so on). (A) The promoter contains a DNA sequence called the TATA box, which is located 25 nucleotides away from the site at which transcription is initiated. (B) The TATA box is recognized and bound by transcription factor TFIID, which then enables the adjacent binding of TFIIB (C). For simplicity the DNA distortion produced by the binding of TFIID (see Figure 6–18) is not shown. (D) The rest of the general transcription factors, as well as the RNA polymerase itself, assemble at the promoter. (E) TFIIH then uses ATP to pry apart the DNA double helix at the transcription start point, allowing transcription to begin. TFIIH also phosphorylates RNA polymerase II, changing its conformation so that the polymerase is released from the general factors and can begin the elongation phase of transcription. As shown, the site of phosphorylation is a long C-terminal polypeptide tail that extends from the polymerase molecule. The assembly scheme shown in the figure was deduced from experiments performed *in vitro*, and the exact order in which the general transcription factors assemble on promoters in cells is not known with certainty. In some cases, the general factors are thought to first assemble with the polymerase, with the whole assembly subsequently binding to the DNA in a single step. The general transcription factors have been highly conserved in evolution; some of those from human cells can be replaced in biochemical experiments by the corresponding factors from simple yeasts.

and listed as TFIIA, TFIIB, and so on. In a broad sense, the eucaryotic general transcription factors carry out functions equivalent to those of the σ factor in bacteria.

Figure 6–16 shows how the general transcription factors assemble *in vitro* at promoters used by RNA polymerase II. The assembly process starts with the binding of the general transcription factor TFIID to a short double-helical DNA sequence primarily composed of T and A nucleotides. For this reason, this sequence is known as the TATA sequence, or **TATA box**, and the subunit of TFIID that recognizes it is called TBP (for TATA-binding protein). The TATA box is typically located 25 nucleotides upstream from the transcription start site. It is not the only DNA sequence that signals the start of transcription (Figure 6–17), but



element	consensus sequence	general transcription factor
BRE	G/C G/C G/A C G C C	TFIIB
TATA	T A T A A/T A A/T	TBP
INR	C/T C/T A N T/A C/T C/T	TFIID
DPE	A/G G A/T C G T G	TFIID

Figure 6–17 Consensus sequences found in the vicinity of eucaryotic RNA polymerase II start points. The name given to each consensus sequence *(first column)* and the general transcription factor that recognizes it *(last column)* are indicated. N indicates any nucleotide, and two nucleotides separated by a slash indicate an equal probability of either nucleotide at the indicated position. In reality, each consensus sequence is a shorthand representation of a histogram similar to that of Figure 6–12. For most RNA polymerase II transcription start points, only two or three of the four sequences are present. For example, most polymerase II promoters have a TATA box sequence, and those that do not typically have a "strong" INR sequence. Although most of the DNA sequences that influence transcription initiation are located "upstream" of the transcription start point, a few, such as the DPE shown in the figure, are located in the transcribed region.

for most polymerase II promoters, it is the most important. The binding of TFIID causes a large distortion in the DNA of the TATA box (Figure 6–18). This distortion is thought to serve as a physical landmark for the location of an active promoter in the midst of a very large genome, and it brings DNA sequences on both sides of the distortion together to allow for subsequent protein assembly steps. Other factors are then assembled, along with RNA polymerase II, to form a complete *transcription initiation complex* (see Figure 6–16).

After RNA polymerase II has been guided onto the promoter DNA to form a transcription initiation complex, it must gain access to the template strand at the transcription start point. This step is aided by one of the general transcription factors, TFIIH, which contains a DNA helicase. Next, like the bacterial polymerase, polymerase II remains at the promoter, synthesizing short lengths of RNA until it undergoes a conformational change and is released to begin transcribing a gene. A key step in this release is the addition of phosphate groups to the "tail" of the RNA polymerase (known as the CTD or C-terminal domain). This phosphorylation is also catalyzed by TFIIH, which, in addition to a helicase, contains a protein kinase as one of its subunits (see Figure 6–16, D and E). The polymerase can then disengage from the cluster of general transcription factors, undergoing a series of conformational changes that tighten its interaction with DNA and acquiring new proteins that allow it to transcribe for long distances without dissociating.

Once the polymerase II has begun elongating the RNA transcript, most of the general transcription factors are released from the DNA so that they are available to initiate another round of transcription with a new RNA polymerase molecule. As we see shortly, the phosphorylation of the tail of RNA polymerase II also causes components of the RNA processing machinery to load onto the polymerase and thus be in position to modify the newly transcribed RNA as it emerges from the polymerase.

Polymerase II Also Requires Activator, Mediator, and Chromatin-modifying Proteins

The model for transcription initiation just described was established by studying the action of RNA polymerase II and its general transcription factors on purified DNA templates *in vitro*. However, as discussed in Chapter 4, DNA in eucaryotic cells is packaged into nucleosomes, which are further arranged in higher-order chromatin structures. As a result, transcription initiation in a eucaryotic cell is more complex and requires more proteins than it does on purified DNA. First, gene regulatory proteins known as *transcriptional activators*



Figure 6–18 Three-dimensional structure of TBP (TATA-binding protein) bound to DNA. The TBP is the subunit of the general transcription factor TFIID that is responsible for recognizing and binding to the TATA box sequence in the DNA *(red)*. The unique DNA bending caused by TBP—two kinks in the double helix separated by partly unwound DNA—may serve as a landmark that helps to attract the other general transcription factors. TBP is a single polypeptide chain that is folded into two very similar domains *(blue* and *green)*. (Adapted from J.L. Kim et al., *Nature* 365:520–527, 1993.)

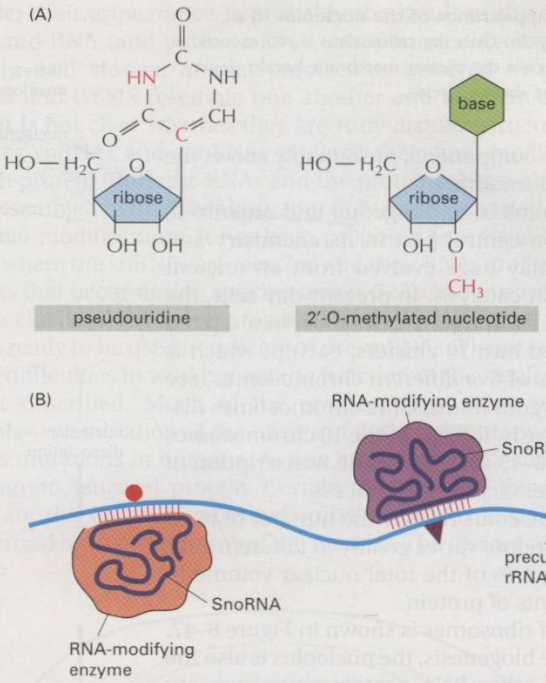


Figure 6–43 Modifications of the precursor rRNA by guide RNAs. (A) Two prominent covalent modifications occur after rRNA synthesis; the differences from the initially incorporated nucleotide are indicated by *red* atoms. (B) As indicated, snoRNAs locate the sites of modification by base-pairing to complementary sequences on the precursor rRNA. The snoRNAs are bound to proteins, and the complexes are called snoRNPs. snoRNPs contain the RNA modification activities, presumably contributed by the proteins but possibly by the snoRNAs themselves.

especially those encoding ribosomal proteins. They are therefore synthesized by RNA polymerase II and processed from excised intron sequences.

The Nucleolus Is a Ribosome-Producing Factory

The nucleolus is the most obvious structure seen in the nucleus of a eucaryotic cell when viewed in the light microscope. Consequently, it was so closely scrutinized by early cytologists that an 1898 review could list some 700 references. We now know that the nucleolus is the site for the processing of rRNAs and their assembly into ribosomes. Unlike other organelles in the cell, it is not bound by a membrane (Figure 6–44); instead, it is a large aggregate of macromolecules, including the rRNA genes themselves, precursor rRNAs, mature rRNAs, rRNA-processing enzymes, snoRNPs, ribosomal protein subunits and partly assembled

Figure 6–44 Electron micrograph of a thin section of a nucleolus in a human fibroblast, showing its three distinct zones. (A) View of entire nucleus. (B) High-power view of the nucleolus. It is believed that transcription of the rRNA genes takes place between the fibrillar center and the dense fibrillar component and that processing of the rRNAs and their assembly into ribosomes proceeds outward from the dense fibrillar component to the surrounding granular components. (Courtesy of E.G. Jordan and J. McGovern.)

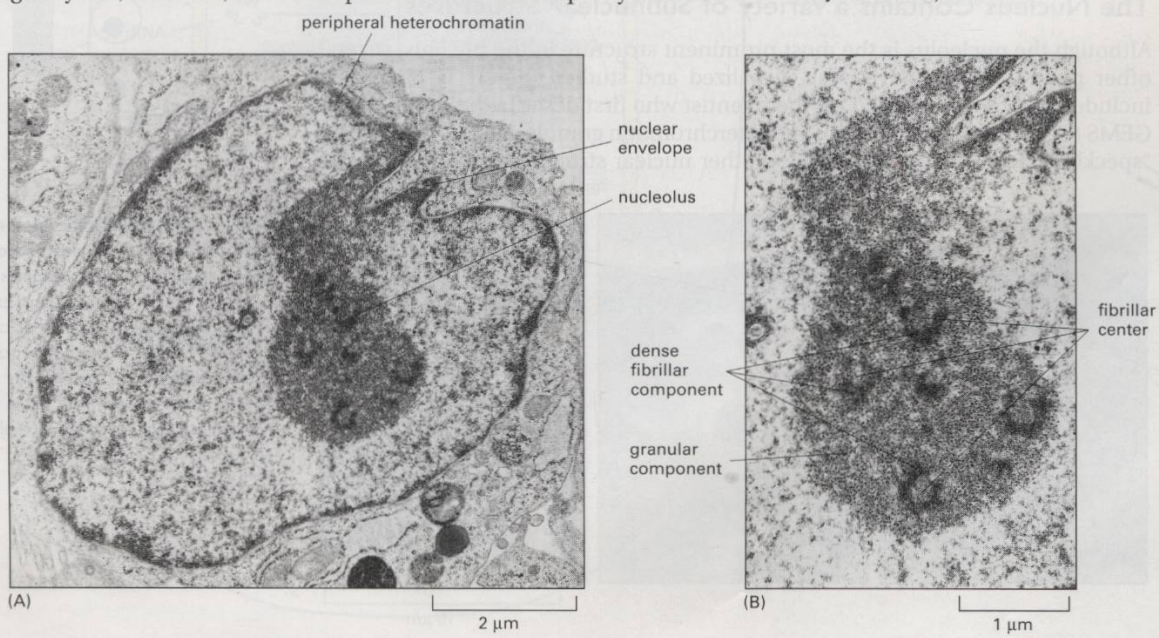
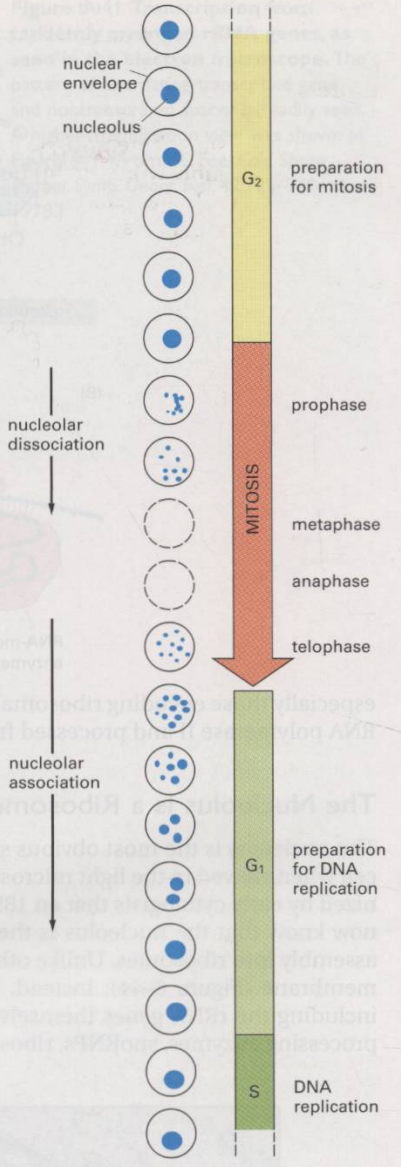


Figure 6–45 Changes in the appearance of the nucleolus in a human cell during the cell cycle. Only the cell nucleus is represented in this diagram. In most eucaryotic cells the nuclear membrane breaks down during mitosis, as indicated by the dashed circles.



ribosomes. The close association of all these components presumably allows the assembly of ribosomes to occur rapidly and smoothly.

It is not yet understood how the nucleolus is held together and organized, but various types of RNA molecules play a central part in its chemistry and structure, suggesting that the nucleolus may have evolved from an ancient structure present in cells dominated by RNA catalysis. In present-day cells, the rRNA genes also have an important role in forming the nucleolus. In a diploid human cell, the rRNA genes are distributed into 10 clusters, each of which is located near the tip of one of the two copies of five different chromosomes (see Figure 4–11). Each time a human cell undergoes mitosis, the chromosomes disperse and the nucleolus breaks up; after mitosis, the tips of the 10 chromosomes coalesce as the nucleolus reforms (Figures 6–45 and 6–46). The transcription of the rRNA genes by RNA polymerase I is necessary for this process.

As might be expected, the size of the nucleolus reflects the number of ribosomes that the cell is producing. Its size therefore varies greatly in different cells and can change in a single cell, occupying 25% of the total nuclear volume in cells that are making unusually large amounts of protein.

A schematic diagram of the assembly of ribosomes is shown in Figure 6–47. In addition to its important role in ribosome biogenesis, the nucleolus is also the site where other RNAs are produced and other RNA–protein complexes are assembled. For example, the U6 snRNP, which, as we have seen, functions in pre-mRNA splicing (see Figure 6–29), is composed of one RNA molecule and at least seven proteins. The U6 snRNA is chemically modified by snoRNAs in the nucleolus before its final assembly there into the U6 snRNP. Other important RNA protein complexes, including telomerase (encountered in Chapter 5) and the signal recognition particle (which we discuss in Chapter 12), are also believed to be assembled at the nucleolus. Finally, the tRNAs (transfer RNAs) that carry the amino acids for protein synthesis are processed there as well. Thus, the nucleolus can be thought of as a large factory at which many different noncoding RNAs are processed and assembled with proteins to form a large variety of ribonucleoprotein complexes.

The Nucleus Contains a Variety of Subnuclear Structures

Although the nucleolus is the most prominent structure in the nucleus, several other nuclear bodies have been visualized and studied (Figure 6–48). These include Cajal bodies (named for the scientist who first described them in 1906), GEMS (Gemini of coiled bodies), and interchromatin granule clusters (also called "speckles"). Like the nucleolus, these other nuclear structures lack membranes

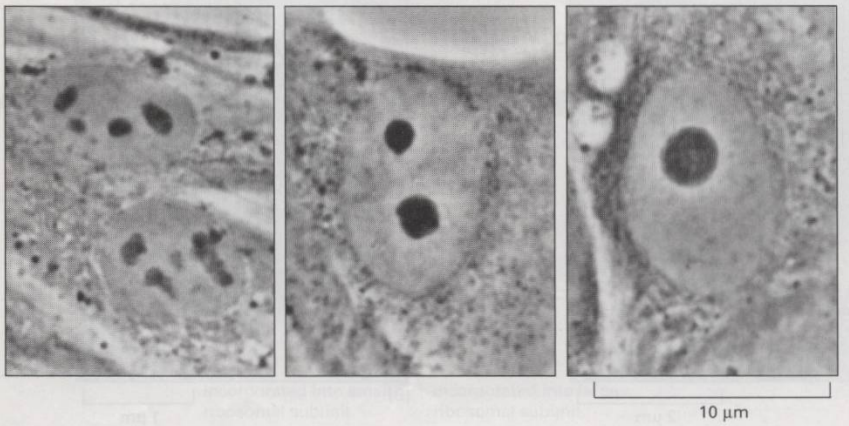


Figure 6–46 Nucleolar fusion. These light micrographs of human fibroblasts grown in culture show various stages of nucleolar fusion. After mitosis, each of the ten human chromosomes that carry a cluster of rRNA genes begins to form a tiny nucleolus, but these rapidly coalesce as they grow to form the single large nucleolus typical of many interphase cells. (Courtesy of E.G. Jordan and J. McGovern.)

and are highly dynamic; their appearance is probably the result of the tight association of protein and RNA (and perhaps DNA) components involved in the synthesis, assembly, and storage of macromolecules involved in gene expression. Cajal bodies and GEMS resemble one another and are frequently paired in the nucleus; it is not clear whether they are truly distinct structures. They may be sites where snRNAs and snoRNAs undergo their final modifications and assembly with protein. Both the RNAs and the proteins that make up the snRNPs are partly assembled in the cytoplasm, but they are transported into the nucleus for their final modifications. It has been proposed that Cajal bodies/GEMS are also sites where the snRNPs are recycled and their RNAs are "reset" after the rearrangements that occur during splicing (see p. 322). In contrast, the interchromatin granule clusters have been proposed to be stockpiles of fully mature snRNPs that are ready to be used in splicing of pre-mRNAs (Figure 6–49).

Scientists have had difficulties in working out the function of the small subnuclear structures just described. Much of the progress now being made depends on genetic tools—examination of the effects of designed mutations in mice or of spontaneous mutations in humans. As one example, GEMS contain the SMN (survival of motor neurons) protein. Certain mutations of the gene encoding this protein are the cause of inherited spinal muscular atrophy, a human disease characterized by a wasting away of the muscles. The disease seems

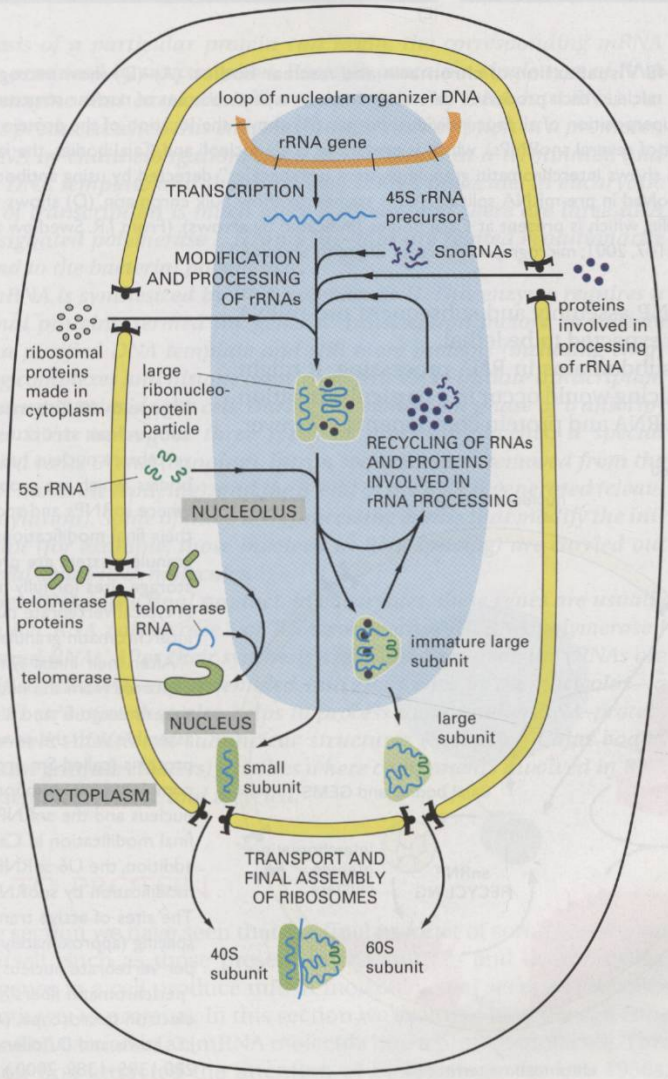


Figure 6–47 The function of the nucleolus in ribosome and other ribonucleoprotein synthesis. The 45S precursor rRNA is packaged in a large ribonucleoprotein particle containing many ribosomal proteins imported from the cytoplasm. While this particle remains in the nucleolus, selected pieces are added and others discarded as it is processed into immature large and small ribosomal subunits. The two ribosomal subunits are thought to attain their final functional form only as each is individually transported through the nuclear pores into the cytoplasm. Other ribonucleoprotein complexes, including telomerase shown here, are also assembled in the nucleolus.

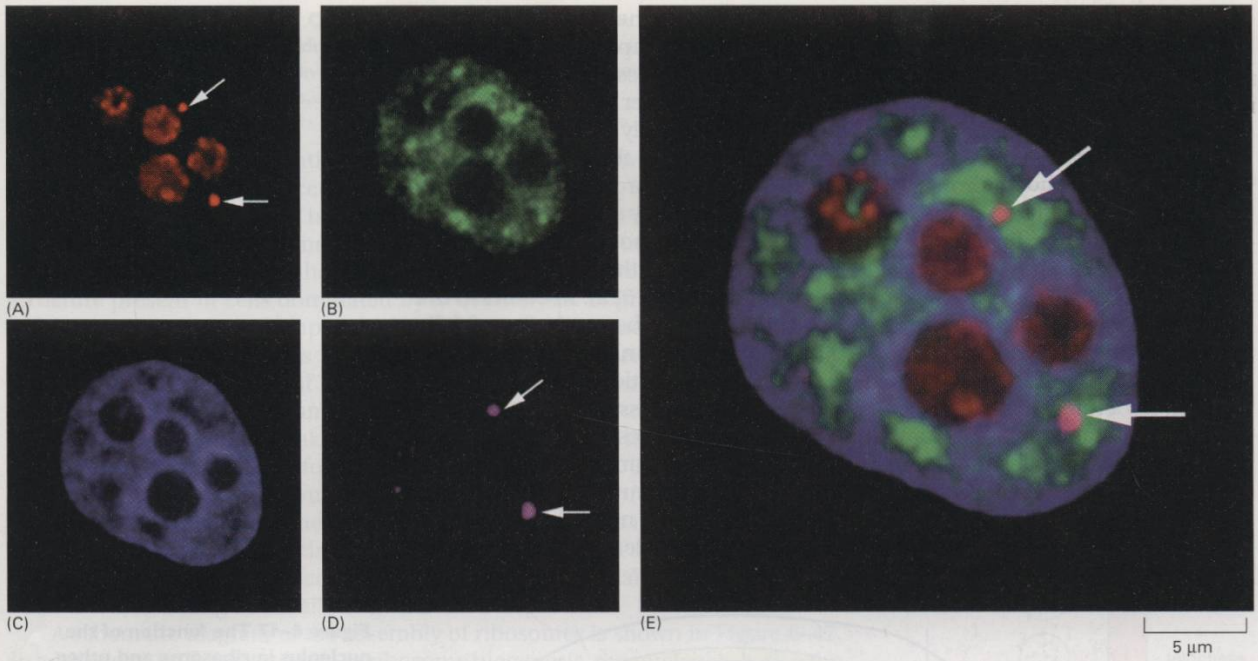


Figure 6–48 Visualization of chromatin and nuclear bodies. (A)–(D) show micrographs of the same human cell nucleus, each processed differently to show a particular set of nuclear structures. (E) shows an enlarged superposition of all four individual images. (A) shows the location of the protein fibrillarin (a component of several snoRNPs), which is present at both nucleoli and Cajal bodies, the latter indicated by arrows. (B) shows interchromatin granule clusters or "speckles" detected by using antibodies against a protein involved in pre-mRNA splicing. (C) is stained to show bulk chromatin. (D) shows the location of the protein coilin, which is present at Cajal bodies (indicated by arrows). (From J.R. Swedlow and A.I. Lamond, *Gen. Biol.* 2:1–7, 2001; micrographs courtesy of Judith Sleeman.)

to be caused by a subtle defect in snRNP assembly and subsequent pre-mRNA splicing. More severe defects would be expected to be lethal.

Given the importance of nuclear subdomains in RNA processing, it might have been expected that pre-mRNA splicing would occur in a particular location in the nucleus, as it requires numerous RNA and protein components. However,

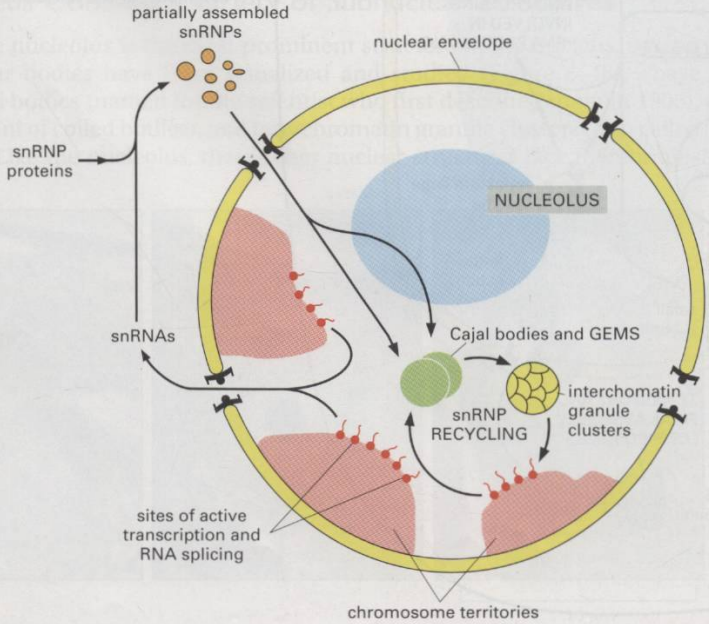


Figure 6–49 Schematic view of subnuclear structures. A typical vertebrate nucleus has several Cajal bodies, which are proposed to be the sites where snRNPs and snoRNPs undergo their final modifications. Interchromatin granule clusters are proposed to be storage sites for fully mature snRNPs. A typical vertebrate nucleus has 20–50 interchromatin granule clusters.

After their initial synthesis, snRNAs are exported from the nucleus, after which they undergo 5′ and 3′ end-processing and assemble with the seven common snRNP proteins (called Sm proteins). These complexes are reimported into the nucleus and the snRNPs undergo their final modification in Cajal bodies. In addition, the U6 snRNP requires chemical modification by snoRNAs in the nucleolus. The sites of active transcription and splicing (approximately 2000–3000 sites per vertebrate nucleus) correspond to the "perichromatin fibers" seen under the electron microscope. (Adapted from J.D. Lewis and D. Tollervey, *Science* 288:1385–1389, 2000.)

we have seen that the assembly of splicing components on pre-mRNA is co-transcriptional; thus splicing must occur at many locations along chromosomes. We saw in Chapter 4 that interphase chromosomes occupy discrete territories in the nucleus, and transcription and pre-mRNA splicing must take place within these territories. However, interphase chromosomes are themselves dynamic and their exact positioning in the nucleus correlates with gene expression. For example, transcriptionally silent regions of interphase chromosomes are often associated with the nuclear envelope where the concentration of heterochromatin components is believed to be especially high. When these same regions become transcriptionally active, they relocate towards the interior of the nucleus, which is richer in the components required for mRNA synthesis. It has been proposed that, although a typical mammalian cell may be expressing on the order of 15,000 genes, transcription and RNA splicing may be localized to only several thousand sites in the nucleus. These sites themselves are highly dynamic and probably result from the association of transcription and splicing components to create small "assembly lines" where the local concentration of these components is very high. As a result, the nucleus seems to be highly organized into subdomains, with snRNPs, snoRNPs, and other nuclear components moving between them in an orderly fashion according to the needs of the cell (Figure 6–49).

Summary

Before the synthesis of a particular protein can begin, the corresponding mRNA molecule must be produced by transcription. Bacteria contain a single type of RNA polymerase (the enzyme that carries out the transcription of DNA into RNA). An mRNA molecule is produced when this enzyme initiates transcription at a promoter, synthesizes the RNA by chain elongation, stops transcription at a terminator, and releases both the DNA template and the completed mRNA molecule. In eucaryotic cells, the process of transcription is much more complex, and there are three RNA polymerases—designated polymerase I, II, and III—that are related evolutionarily to one another and to the bacterial polymerase.

Eucaryotic mRNA is synthesized by RNA polymerase II. This enzyme requires a series of additional proteins, termed the general transcription factors, to initiate transcription on a purified DNA template and still more proteins (including chromatin-remodeling complexes and histone acetyltransferases) to initiate transcription on its chromatin template inside the cell. During the elongation phase of transcription, the nascent RNA undergoes three types of processing events: a special nucleotide is added to its 5′ end (capping), intron sequences are removed from the middle of the RNA molecule (splicing), and the 3′ end of the RNA is generated (cleavage and polyadenylation). Some of these RNA processing events that modify the initial RNA transcript (for example, those involved in RNA splicing) are carried out primarily by special small RNA molecules.

For some genes, RNA is the final product. In eucaryotes, these genes are usually transcribed by either RNA polymerase I or RNA polymerase III. RNA polymerase I makes the ribosomal RNAs. After their synthesis as a large precursor, the rRNAs are chemically modified, cleaved, and assembled into ribosomes in the nucleolus—a distinct subnuclear structure that also helps to process some smaller RNA–protein complexes in the cell. Additional subnuclear structures (including Cajal bodies and interchromatin granule clusters) are sites where components involved in RNA processing are assembled, stored, and recycled.

FROM RNA TO PROTEIN

In the preceding section we have seen that the final product of some genes is an RNA molecule itself, such as those present in the snRNPs and in ribosomes. However, most genes in a cell produce mRNA molecules that serve as intermediaries on the pathway to proteins. In this section we examine how the cell converts the information carried in an mRNA molecule into a protein molecule. This feat of translation first attracted the attention of biologists in the late 1950s,

GCA GCC GCG GCU	AGA AGG CGA CGC CGG CGU	GAC GAU	AAC AAU	UGC UGU	GAA GAG	CAA CAG	GGA GGC GGG GGU	CAC CAU	AUA AUC AUU	UUA UUG CUA CUC CUG CUU	AAA AAG	AUG	UUC UUU	CCA CCC CCG CCU	AGC AGU UCA UCC UCG UCU	ACA ACC ACG ACU	UGG	UAC UAU	GUA GUC GUG GUU	UAA UAG UGA
Ala	Arg	Asp	Asn	Cys	Glu	Gln	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	stop
A	R	D	N	C	E	Q	G	H	I	L	K	M	F	P	S	T	W	Y	V	

Figure 6–50 The genetic code. The standard one-letter abbreviation for each amino acid is presented below its three-letter abbreviation (see Panel 3–1, pp. 132–133, for the full name of each amino acid and its structure). By convention, codons are always written with the 5′-terminal nucleotide to the left. Note that most amino acids are represented by more than one codon, and that there are some regularities in the set of codons that specifies each amino acid. Codons for the same amino acid tend to contain the same nucleotides at the first and second positions, and vary at the third position. Three codons do not specify any amino acid but act as termination sites (stop codons), signaling the end of the protein-coding sequence. One codon—AUG—acts both as an initiation codon, signaling the start of a protein-coding message, and also as the codon that specifies methionine.

when it was posed as the "coding problem": how is the information in a linear sequence of nucleotides in RNA translated into the linear sequence of a chemically quite different set of subunits—the amino acids in proteins? This fascinating question stimulated great excitement among scientists at the time. Here was a cryptogram set up by nature that, after more than 3 billion years of evolution, could finally be solved by one of the products of evolution—human beings. And indeed, not only has the code been cracked step by step, but in the year 2000 the elaborate machinery by which cells read this code—the ribosome—was finally revealed in atomic detail.

An mRNA Sequence Is Decoded in Sets of Three Nucleotides

Once an mRNA has been produced, by transcription and processing the information present in its nucleotide sequence is used to synthesize a protein. Transcription is simple to understand as a means of information transfer: since DNA and RNA are chemically and structurally similar, the DNA can act as a direct template for the synthesis of RNA by complementary base-pairing. As the term *transcription* signifies, it is as if a message written out by hand is being converted, say, into a typewritten text. The language itself and the form of the message do not change, and the symbols used are closely related.

In contrast, the conversion of the information in RNA into protein represents a **translation** of the information into another language that uses quite different symbols. Moreover, since there are only four different nucleotides in mRNA and twenty different types of amino acids in a protein, this translation cannot be accounted for by a direct one-to-one correspondence between a nucleotide in RNA and an amino acid in protein. The nucleotide sequence of a gene, through the medium of mRNA, is translated into the amino acid sequence of a protein by rules that are known as the **genetic code**. This code was deciphered in the early 1960s.

The sequence of nucleotides in the mRNA molecule is read consecutively in groups of three. RNA is a linear polymer of four different nucleotides, so there are $4 \times 4 \times 4 = 64$ possible combinations of three nucleotides: the triplets AAA, AUA, AUG, and so on. However, only 20 different amino acids are commonly found in proteins. Either some nucleotide triplets are never used, or the code is redundant and some amino acids are specified by more than one triplet. The second possibility is, in fact, the correct one, as shown by the completely deciphered genetic code in Figure 6–50. Each group of three consecutive nucleotides in RNA is called a **codon**, and each codon specifies either one amino acid or a stop to the translation process.

This genetic code is used universally in all present-day organisms. Although a few slight differences in the code have been found, these are chiefly in the DNA of mitochondria. Mitochondria have their own transcription and protein synthesis systems that operate quite independently from those of the rest of the cell, and it is understandable that their small genomes have been able to accommodate minor changes to the code (discussed in Chapter 14).

In principle, an RNA sequence can be translated in any one of three different **reading frames**, depending on where the decoding process begins (Figure 6–51). However, only one of the three possible reading frames in an mRNA encodes the required protein. We see later how a special punctuation signal at the beginning of each RNA message sets the correct reading frame at the start of protein synthesis.

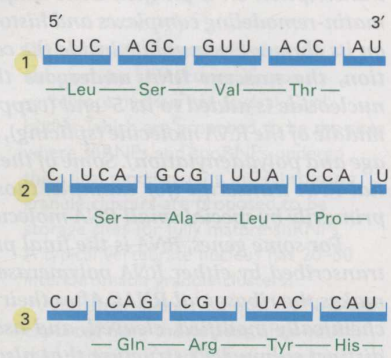


Figure 6–51 The three possible reading frames in protein synthesis. In the process of translating a nucleotide sequence *(blue)* into an amino acid sequence *(green)*, the sequence of nucleotides in an mRNA molecule is read from the 5′ to the 3′ end in sequential sets of three nucleotides. In principle, therefore, the same RNA sequence can specify three completely different amino acid sequences, depending on the reading frame. In reality, however, only one of these reading frames contains the actual message.

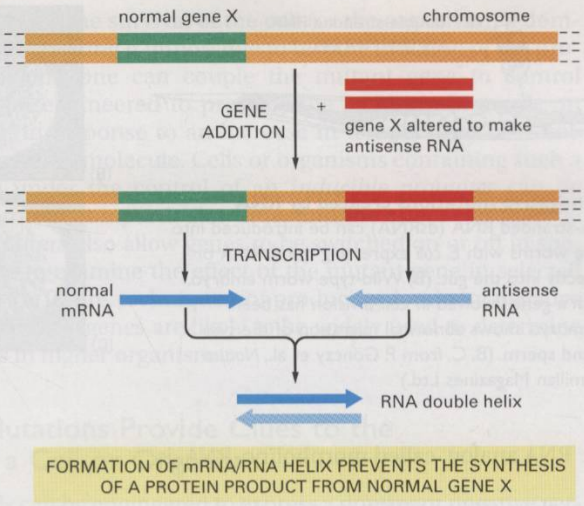


Figure 8–65 The antisense RNA strategy for generating dominant negative mutations. Mutant genes that have been engineered to produce antisense RNA, which is complementary in sequence to the RNA made by the normal gene X, can cause double-stranded RNA to form inside cells. If a large excess of the antisense RNA is produced, it can hybridize with—and thereby inactivate—most of the normal RNA produced by gene X. Although in the future it may become possible to inactivate any gene in this way, at present the technique seems to work for some genes but not others.

this way cells can be made to order that produce an altered form of any specific protein or RNA molecule instead of the normal form of the molecule. If the mutant gene is completely inactive and the gene product normally performs an essential function, the cell dies; but in this case a less severely mutated version of the gene can be used to replace the normal gene, so that the mutant cell survives but is abnormal in the process for which the gene is required. Often the mutant of choice is one that produces a temperature-sensitive gene product, which functions normally at one temperature but is inactivated when cells are shifted to a higher or lower temperature.

The ability to perform direct gene replacements in lower eucaryotes, combined with the power of standard genetic analyses in these haploid organisms, explains in large part why studies in these types of cells have been so important for working out the details of those processes that are shared by all eucaryotes. As we shall see, gene replacements are possible, but more difficult to perform in higher eucaryotes, for reasons that are not entirely understood.

Engineered Genes Can Be Used to Create Specific Dominant Negative Mutations in Diploid Organisms

Higher eucaryotes, such as mammals, fruit flies, or worms, are diploid and therefore have two copies of each chromosome. Moreover, transfection with an altered gene generally leads to gene addition rather than gene replacement: the altered gene inserts at a random location in the genome, so that the cell (or the organism) ends up with the mutated gene in addition to its normal gene copies.

Because gene addition is much more easily accomplished than gene replacement in higher eucaryotic cells, it is useful to create specific dominant negative mutations in which a mutant gene eliminates the activity of its normal counterparts in the cell. One ingenious approach exploits the specificity of hybridization reactions between two complementary nucleic acid chains. Normally, only one of the two DNA strands in a given portion of double helix is transcribed into RNA, and it is always the same strand for a given gene (see Figure 6–14). If a cloned gene is engineered so that the opposite DNA strand is transcribed instead, it will produce antisense RNA molecules that have a sequence complementary to the normal RNA transcripts. Such *antisense RNA*, when synthesized in large enough amounts, can often hybridize with the "sense" RNA made by the normal genes and thereby inhibit the synthesis of the corresponding protein (Figure 8–65). A related method involves synthesizing short antisense nucleic acid molecules chemically or enzymatically and then injecting (or otherwise delivering) them into cells, again blocking (although only temporarily) production of the corresponding protein. To avoid degradation of the

This material may be protected by Copyright law (Title 17 U.S. Code)

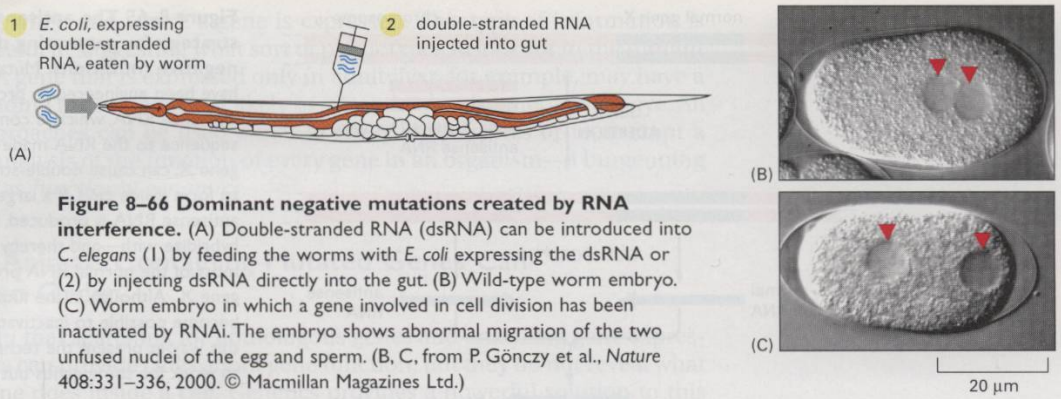


Figure 8–66 Dominant negative mutations created by RNA interference. (A) Double-stranded RNA (dsRNA) can be introduced into *C. elegans* (1) by feeding the worms with *E. coli* expressing the dsRNA or (2) by injecting dsRNA directly into the gut. (B) Wild-type worm embryo. (C) Worm embryo in which a gene involved in cell division has been inactivated by RNAi. The embryo shows abnormal migration of the two unfused nuclei of the egg and sperm. (B, C, from P. Gönczy et al., *Nature* 408:331–336, 2000. © Macmillan Magazines Ltd.)

injected nucleic acid, a stable synthetic RNA analog, called morpholino-RNA, is often used instead of ordinary RNA.

As investigators continued to explore the antisense RNA strategy, they made an interesting discovery. An antisense RNA strand can block gene expression, but a preparation of double-stranded RNA (dsRNA), containing both the sense and antisense strands of a target gene, inhibit the activity of target genes even more effectively (see Figure 7–107). This phenomenon, dubbed *RNA interference (RNAi),* has now been exploited for examining gene function in several organisms.

The RNAi technique has been widely used to study gene function in the nematode *C. elegans*. When working with worms, introducing the dsRNA is quite simple: RNA can be injected directly into the intestine of the animal, or the worm can be fed with *E. coli* expressing the target gene dsRNA (Figure 8–66A). The RNA is distributed throughout the body of the worm and is found to inhibit expression of the target gene in different tissue types. Further, as explained in Figure 7–107, the interference is frequently inherited by the progeny of the injected animal. Because the entire genome of *C. elegans* has been sequenced, RNAi is being used to help in assigning functions to the entire complement of worm genes. In one study, researchers were able to inhibit 96% of the approximately 2300 predicted genes on *C. elegans* chromosome III. In this way, they identified 133 genes involved in cell division in *C. elegans* embryos (Figure 8–66C). Of these, only 11 had been previously ascribed a function by direct experimentation.

For unknown reasons, RNA interference does not efficiently inactivate all genes. And interference can sometimes suppress the activity of a target gene in one tissue and not another. An alternative way to produce a dominant negative mutation takes advantage of the fact that most proteins function as part of a larger protein complex. Such complexes can often be inactivated by the inclusion of just one nonfunctional component. Therefore, by designing a gene that produces large quantities of a mutant protein that is inactive but still able to assemble into the complex, it is often possible to produce a cell in which all the complexes are inactivated despite the presence of the normal protein (Figure 8–67).

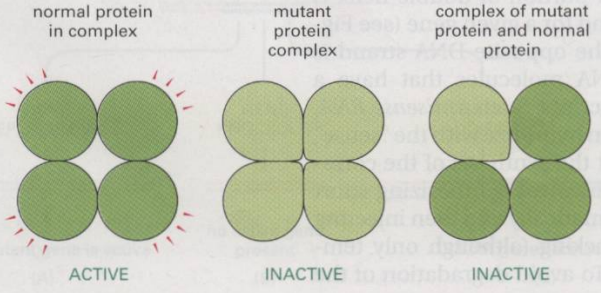


Figure 8–67 A dominant negative effect of a protein. Here a gene is engineered to produce a mutant protein that prevents the normal copies of the same protein from performing their function. In this simple example, the normal protein must form a multisubunit complex to be active, and the mutant protein blocks function by forming a mixed complex that is inactive. In this way a single copy of a mutant gene located anywhere in the genome can inactivate the normal products produced by other gene copies.

If a protein is required for the survival of the cell (or the organism), a dominant negative mutant dies, making it impossible to test the function of the protein. To avoid this problem, one can couple the mutant gene to control sequences that have been engineered to produce the gene product only on command—for example, in response to an increase in temperature or to the presence of a specific signaling molecule. Cells or organisms containing such a dominant mutant gene under the control of an *inducible promoter* can be deprived of a specific protein at a particular time, and the effect can then be followed. Inducible promoters also allow genes to be switched on or off in specific tissues, allowing one to examine the effect of the mutant gene in selected parts of the organism. In the future, techniques for producing dominant negative mutations to inactivate specific genes are likely to be widely used to determine the functions of proteins in higher organisms.

Gain-of-Function Mutations Provide Clues to the Role Genes Play in a Cell or Organism

In the same way that cells can be engineered to express a dominant negative version of a protein, resulting in a loss-of-function phenotype, they can also be engineered to display a novel phenotype through a *gain-of-function* mutation. Such mutations may confer a novel activity on a particular protein, or they may cause a protein with normal activity to be expressed at an inappropriate time or in the wrong tissue in an animal. Regardless of the mechanism, gain-of-function mutations can produce a new phenotype in a cell, tissue, or organism.

Often, gain-of-function mutants are generated by expressing a gene at a much higher level than normal in cells. Such overexpression can be achieved by coupling a gene to a powerful promoter sequence and placing it on a multicopy plasmid—or integrating it in multiple copies in the genome. In either case, the gene is present in many copies and each copy directs the transcription of unusually large numbers of mRNA molecules. Although the effect that such overexpression has on the phenotype of an organism must be interpreted with caution, this approach has provided invaluable insights into the activity of many genes. In an alternate type of gain-of-function mutation, the mutant protein is made in normal amounts, but is much more active than its normal counterpart. Such proteins are frequently found in tumors, and they have been exploited to study signal transduction pathways in cells (discussed in Chapter 15).

Genes can also be expressed at the wrong time or in the wrong place in an organism—often with striking results (Figure 8–68). Such misexpression is most often accomplished by re-engineering the genes themselves, thereby supplying them with the regulatory sequences needed to alter their expression.

Genes Can Be Redesigned to Produce Proteins of Any Desired Sequence

In studying the action of a gene and the protein it encodes, one does not always wish to make drastic changes—flooding cells with huge quantities of hyperactive protein or eliminating a gene product entirely. It is sometimes useful to make slight changes in a protein's structure so that one can begin to dissect which portions of a protein are important for its function. The activity of an enzyme, for example, can be studied by changing a single amino acid in its active site. Special techniques are required to alter genes, and their protein products, in such subtle ways. The first step is often the chemical synthesis of a short DNA molecule containing the desired altered portion of the gene's nucleotide sequence. This synthetic DNA oligonucleotide is hybridized with single-stranded plasmid DNA that contains the DNA sequence to be altered, using conditions that allow imperfectly matched DNA strands to pair (Figure 8–69). The synthetic oligonucleotide will now serve as a primer for DNA synthesis by DNA polymerase, thereby generating a DNA double helix that incorporates the altered sequence into one of its two strands. After transfection, plasmids that carry the fully modified gene sequence are obtained. The appropriate DNA is



Figure 8–68 Ectopic misexpression of Wnt, a signaling protein that affects development of the body axis in the early *Xenopus* embryo. In this experiment, mRNA coding for Wnt was injected into the ventral vegetal blastomere, inducing a second body axis (discussed in Chapter 21). (From S. Sokol et al., *Cell* 67:741–752, 1992. © Elsevier.)

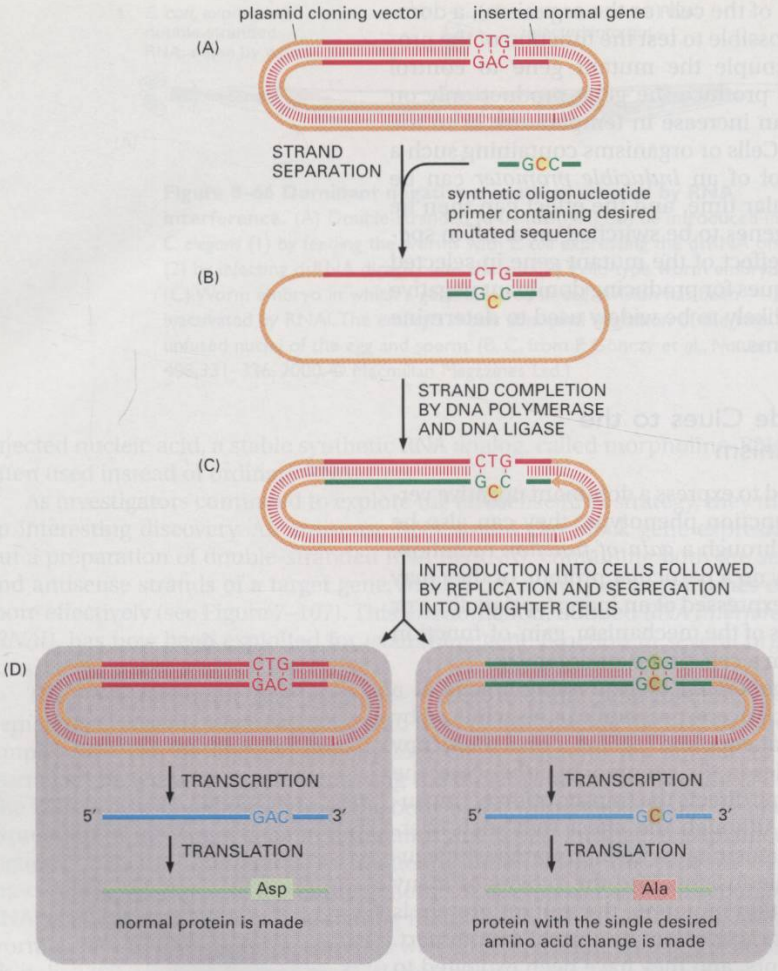


Figure 8–69 The use of a synthetic oligonucleotide to modify the protein-coding region of a gene by site-directed mutagenesis. (A) A recombinant plasmid containing a gene insert is separated into its two DNA strands. A synthetic oligonucleotide primer corresponding to part of the gene sequence but containing a single altered nucleotide at a predetermined point is added to the single-stranded DNA under conditions that permit less than perfect DNA hybridization. (B) The primer hybridizes to the DNA, forming a single mismatched nucleotide pair. (C) The recombinant plasmid is made double-stranded by *in vitro* DNA synthesis starting from the primer and sealed by DNA ligase. (D) The double-stranded DNA is introduced into a cell, where it is replicated. Replication using one strand of the template produces a normal DNA molecule, but replication using the other (the strand that contains the primer) produces a DNA molecule carrying the desired mutation. Only half of the progeny cells will end up with a plasmid that contains the desired mutant gene. However, a progeny cell that contains the mutated gene can be identified, separated from other cells, and cultured to produce a pure population of cells, all of which carry the mutated gene. Only one of the many changes that can be engineered in this way is shown here. With an oligonucleotide of the appropriate sequence, more than one amino acid substitution can be made at a time, or one or more amino acids can be inserted or deleted. Although not shown in this figure, it is also possible to create a site-directed mutation by using the appropriate oligonucleotides and PCR (instead of plasmid replication) to amplify the mutated gene.

then inserted into an expression vector so that the redesigned protein can be produced in the appropriate type of cells for detailed studies of its function. By changing selected amino acids in a protein in this way—a technique called **site-directed mutagenesis**—one can determine exactly which parts of the polypeptide chain are important for such processes as protein folding, interactions with other proteins, and enzymatic catalysis.

Engineered Genes Can Be Easily Inserted into the Germ Line of Many Animals

When engineering an organism that is to express an altered gene, ideally one would like to be able to replace the normal gene with the altered one so that the function of the mutant protein can be analyzed in the absence of the normal protein. As discussed above, this can be readily accomplished in some haploid, single-celled organisms. We shall see in the following section that much more complicated procedures have been developed that allow gene replacements of this type in mice. Foreign DNA can, however, be rather easily integrated into random positions of many animal genomes. In mammals, for example, linear DNA fragments introduced into cells are rapidly ligated end-to-end by intracellular enzymes to form long tandem arrays, which usually become integrated into a chromosome at an apparently random site. Fertilized mammalian eggs behave like other mammalian cells in this respect. A mouse egg injected with 200 copies of a linear DNA molecule often develops into a mouse containing, in many of its cells, a tandem array of copies of the injected gene integrated at a single random