Multifunctional Proteins

Edited by

Hans Bisswanger and Eva Schmincke-Ott
Physiologisch-chemisches Institut
der Universität Tübingen, Germany

A WILEY-INTERSCIENCE PUBLICATION

JOHN WILEY & SONS, New York • Chichester • Brisbane • Toronto

Library of Congress Cataloging in Publication:

Main entry under title:
Multifunctional proteins.

"A Wiley-Interscience publication."
Includes index.
1. Proteins. 2. Enzymes. 3. Structure-activity relationship (Pharmacology) I. Bisswanger, Hans. II. Schmincke-Ott, Eva. [DNLM: 1. Proteins—Physiology. QU55.3 M961]

QP551.M84 574.1'9245 79-16055
ISBN 0-471-04270-6

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

CONTRIBUTORS

Hans Bisswanger, Physiologisch-chemisches Institut der Universität Tübingen, Tübingen, West Germany

Dietmar G. Braun, Basel Institute for Immunology, Basel, Switzerland

Georges N. Cohen, Unité de Biochimie Cellulaire, Département de Biochimie et Génétique Microbienne, Institut Pasteur, Paris, France

R. John Collier, Department of Microbiology, College of Letters and Science and The Molecular Biology Institute, University of California, Los Angeles, California

Irving P. Crawford, Department of Microbiology, University of Iowa, Iowa City, Iowa

Norman P. Curthoys, Department of Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania

Alice Dautry-Varsat, Unité de Biochimie Cellulaire, Département de Biochimie et Génétique Microbienne, Institut Pasteur, Paris, France

Rebecca P. Hughey, Department of Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania

Jürgen Kania, Universitätskliniken Köln, Abteilung Experimentelle Innere Medizin, Köln, West Germany

Horst Kleinkauf, Max-Volmer-Institut für Physikalische Chemie und Molekularbiologie, Abteilung Biochemie, Technische Universität Berlin, Berlin, West Germany

Hans Klenow, Biokemisk Institut B, Panum Instituttet, University of Copenhagen, Denmark

Hans Koischwitz, Max-Volmer-Institut für Physikalische Chemie und Molekularbiologie, Abteilung Biochemie, Technische Universität Berlin, Berlin, West Germany

John J. Mekalanos, Department of Microbiology, College of Letters and Science, University of California, Los Angeles, California

Benno Müller-Hill, Institut für Genetik der Universität Köln, Köln, West Germany

Wolfgang Schalch, Basel Institute for Immunology, Basel, Switzerland

Eva Schmincke-Ott, Physiologisch-chemisches Institut der Universität Tübingen, Tübingen, West Germany

Eckhart Schweizer, Lehrstuhl für Biochemie Institut für Mikrobiologie und Biochemie Universität Erlangen-Nürnberg, Erlangen, West Germany

Howard Zalkin, Department of Biochemistry, Purdue University, West Lafayette, Indiana

PREFACE

A few years ago Kaspar Kirschner and I decided to write a review article for *Annual Reviews of Biochemistry* (K. Kirschner and H. Bisswanger, *Annual Reviews of Biochemistry*, 1977). Our interest was stimulated by the work on the multifunctional protein phosphoribosyl-anthranilate isomerase: indoleglycerol-phosphate synthase, an enzyme that carries two different catalytic functions on one polypeptide chain. We wondered whether proteins of this type originated only from random fusion of two adjacent structural genes and rather expected to find for our review a collection of peculiar proteins whose structure would merely reflect a whim of nature. As the search went on the number of examples grew, and we finally had to accept the fact that the structural principle of multifunctional proteins was that of a class of proteins comparable to multienzyme complexes. It proved to be impossible to cover all aspects in one article, and we had to be content with a general view of the field.

The favorable reaction to this review suggested that, in the meantime, multifunctional proteins have been recognized as an interesting and relevant class of proteins by many biochemists.

When I shifted my field of investigation, working now with Eva Schmincke-Ott on multienzyme complexes, the problem of multifunctionality remained of interest to us, and we decided to start this book in order to have a more thorough and expert description of multifunctional proteins than the previous review could afford.

It is our aim to demonstrate a series of individual systems and by way of their detailed characterization point to different typical aspects of multifunctional proteins. The introduction summarizes these aspects and dwells on some interesting questions raised by the evolution of multifunctional proteins. It also gives a compilation of the known and reasonably well investigated examples to date.

It was rather difficult to find an exact definition of multifunctional proteins since most proteins will comprise several functions depending on how one would conceive the term "function." Allosteric proteins for instance may unite a catalytic and a regulatory function on one polypeptide chain,

proteinases and other proenzymes carry a covalently bound protective peptide, and a large number of proteins combine binding or structural functions with some biological activity.

On the other hand, there are indeed many obviously multifunctional proteins, and in the definition of multifunctional proteins we used we referred to proteins that carry on one polypeptide chain two or more virtually autonomous functions that in principle are measurable or occur independently of one another. (The latter is sometimes found when an organism other than the one producing the multifunctional protein is investigated.) Of course there are always examples that escape a clear-cut definition, therefore to some extent our selection is but a personal judgment, and the table of multifunctional proteins in the introduction must be read in a qualified sense, keeping in mind the difficulties that arise when defining independent functions or interpreting yet incomplete experimental evidence in order to achieve a fairly systematic compilation.

We are aware that this monograph represents only the momentary situation of research in this field. Considering the pace at which protein isolation methods are becoming more effective and the increased availability of protein characterizations, it is difficult to cover the continually growing number of multifunctional proteins. However, we hope to have collected enough examples to convey a useful concept of this field.

We are indebted to several people who have supported our work on this book. They are Dr. V. Braun, Dr. I. P. Crawford, Dr. U. Henning, Dr. K. Kirschner, Dr. E. Schweizer, and Dr. H. Zalkin, who have kindly read the introduction, Dr. R. Clark, who helped with the English phrasing, and Frau G. Knödler, who handled the mail and manuscripts.

HANS BISSWANGER
EVA SCHMINCKE-OTT

Tübingen, Germany
November 1979

CONTENTS

Chapter 1

INTRODUCTION

EVA SCHMINCKE-OTT AND HANS BISSWANGER
Physiologisch-chemisches Institut der Universität Tübingen
Tübingen, Germany

1.1 PERSPECTIVES

Former concepts of the relationship between the structure and function of proteins were based on the hypothesis that one polypeptide chain would be responsible for a single function, and at first the general experimental evidence supported this notion. This view has had to be altered, however, as more refined techniques, especially x-ray diffraction analysis, became available and allowed a more exact and detailed analysis of protein tertiary structure.

Largely as the result of comparative structural studies of dehydrogenases in the laboratory of M. G. Rossman (1, 2), as well as the elucidation of immunoglobulin structures by G. M. Edelman and colaborators (3, 4), it seems reasonable to consider a single polypeptide chain as subdivided into a number of globular regions called domains, each of which is responsible for a certain defined function. Domains then appear to constitute the structural and functional building blocks of proteins.

It seems to be a general principle in biology that complex organizational forms develop as the result of the repetition of simpler structures. An initial

simple multiplication of a number of structural elements is followed by their variegation and specialization such that their common origin is no longer easily recognizable. This principle also appears to hold for the evolution of complex protein systems. Figure 1.1 gives an idea of the possible organizational schemes for complex proteins. In such proteins the combined operation of several structural and functional elements is the rule. The number of possible combinations of such elements is small. Either homogeneous or hetereogeneous elements may be combined, and these may either be covalently bound, or noncovalently associated.

A large number of proteins are composed of subunits that are identical. Most enzymes are constructed in this way (5), and of these some are subject to allosteric regulation. If the identical subunits are covalently connected, a protein containing repeated amino acid sequences results. Examples of this latter class are not rare, but in only a few cases is the original function maintained in all of the repeated sequences.

A smaller proportion of proteins [about 10%—roughly estimated from a compilation of oligomeric proteins by I. Klotz (5)] is made up of heterogeneous subunits. Where these are noncovalently associated they form a multienzyme complex in which autonomous functions are performed by separate polypeptide chains. If, on the other hand, the heterogeneous subunits are covalently connected, a multifunctional protein results in which autonomous functions are distributed among the domains of a polypeptide chain. The number of proteins made up of heterogeneous subunits appears to be about the same in both classes.

However, in comparison with the other classes of proteins multifunctional proteins *per se* have been given little attention until recently. This stems in

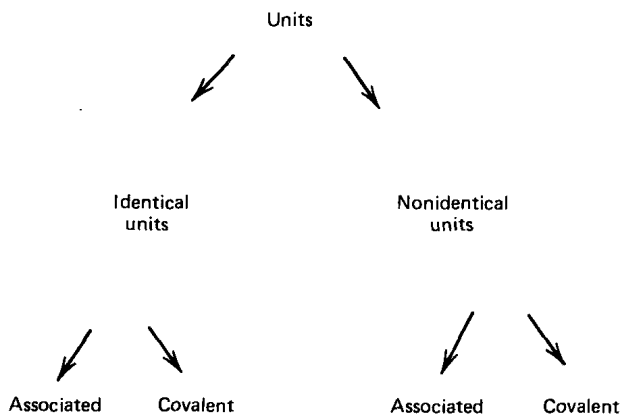

Fig. 1.1 Organizational scheme of complex proteins.

part from the fact that complex proteins with different functions used to be simply regarded as multienzyme complexes. Fatty acid synthetase, for example, used to be taken as an ideal example of a multienzyme complex (6, 7), but more critical investigations have revealed that it is in fact composed of multifunctional polypeptide chains (8–16).

Two simple techniques have led to the discovery of an increasing number of multifunctional proteins: improved purification procedures and polyacrylamide gel electrophoresis under protein denaturing conditions. Purification was improved by using gentle extraction and separation methods and performing them in the presence of proteolysis inhibitors. This is especially important since it has been found that many multifunctional proteins are particularly vulnerable to proteolytic attack in the area between two functional regions. This could very well lead to the destruction of the structural entity while the respective constituent activities remained unimpaired. For instance cytochrome b_5 was isolated, crystallized (17), and submitted to x-ray diffraction studies (18) before it was realized that a proteolytic artefact had been investigated (19). Polyacrylamide gel electrophoresis performed under denaturing conditions has provided a tool for estimating in a fast, simple, and fairly accurate fashion the minimal size of protein subunits.

The growing number of multifunctional proteins (20, 21, 22) shows that we are not simply dealing with a few cases of accidentally fused polypeptide chains, but with a typical organizational scheme for proteins comparable to that of multienzyme complexes. It would be interesting to find out whether the differences between the two organizational schemes are sufficiently significant to explain why in one organism multifunctional proteins occur and in the other multienzyme complexes, and why in some organisms the functions are completely separate.

1.2 CHARACTERIZATION OF MULTIFUNCTIONAL SYSTEMS

Definition of Multifunctional Proteins

Only proteins that combine several autonomous functions on one polypeptide chain should be accepted as multifunctional proteins. In its functional aspect this definition should describe autonomously measurable functions, not ruling out certain regulatory interdependencies between these functions. The definition is also meant to encompass the structural aspect of multifunctionality. Autonomy in this sense implies that each function is assigned to a distinct region, that is a domain, on the polypeptide chain.

Excluded are enzymes that can catalyze different reactions using the same reaction center, like asparaginase which can work as glutaminase (23), glu-

cose-6-phosphatase which can catalyse two reactions (24), or phosphoglyceromutase which can catalyse three reactions always using the same reaction center (25, 26). In these cases we are only dealing with variations of one common function, so that the expression "multifunctional" is not appropriate. Allosteric enzymes and proenzymes will not be considered multifunctional either, although often they are proteins that carry a catalytic function and a regulatory or protective function on different regions of one polypeptide chain. However, these supplementary functions are not autonomous, they are only defined with respect to the catalytic function.

As constituents of multifunctional proteins we find mainly catalytical functions but also binding functions, as long as they are autonomous, like the membrane binding of amphipathic proteins, antigen binding of antibodies, or the DNA binding of lac repressor.

Review of Multifunctional Protein Systems

Table 1.1 includes a collection of the known examples of multifunctional proteins. For most of the systems, there is ample evidence showing them to be multifunctional proteins, but there are some systems listed that lack decisive evidence in this sense. Usually some details of their structure are not known. They are included for the sake of completeness however, and because they will be used as specific examples. The table shows that the class of multifunctional proteins is represented in all organisms from bacteria to mammals.

By far the largest group consists of enzymes that combine two or more functions. There are virtually no examples where the different functions are completely unrelated to each other. Usually they are catalytic functions in the same metabolic pathway, sometimes even of sequential steps. The noncatalytic functions in multifunctional proteins encompass transport and storage functions of yeast hemoglobin (134) and bovine serum albumin (119), the membrane binding of toxins (105) and amphipathic enzymes (94), as well as the binding functions of antibodies (127) and *lac* repressor (122, 123). The latter is an interesting and structurally well investigated example of cooperation between domains, the operator, and inducer binding sites. In myosin the structuring function of light meromyosin that builds up the strong filaments is combined with the function of motility executed by the S1 domain of heavy meromyosin and its ATPase activity (104).

Evidence for the Structure of Multifunctional Proteins

For a protein to qualify as multifunctional it must have had a thorough analysis of its structure. Two important criteria have to be fulfilled:

1 The existence of more than one function on one single polypeptide chain.
2 The autonomy of these functions, which may be demonstrated by the existence of distinct domains for the different functions on this polypeptide chain.

Genetic analysis may produce evidence for the existence of only one polypeptide chain when it is possible to obtain single point mutants deficient in more than one function. Pleiotropic effects must be excluded. If it is impossible to obtain complementation in mutants that are defective in alternative functions this fact may suggest multifunctionality in a protein as well.

The proper physical methods would be ultracentrifugation, gelfiltration, and polyacrylamide gel electrophoresis in the presence of and without denaturing agents. In the case of a multifunctional protein the number of autonomous functions must exceed the number of separable protein bands. One must be aware of the homogenity of the protein band, this may be achieved by endgroup analysis or the comparison of the theoretical number of tryptic peptides with the experimental amount.

Genetic analysis may also indicate autonomous domains if it is possible to isolate mutants that are defective in only one function. At the expense of time and means, very conclusive evidence may be derived from setting up an exact genetic map as was done for phosphoribosyl-anthranilate isomerase: indoleglycerolphosphate synthase from *E. coli* (37, 135). By a thorough analysis of a large collection of mutations, illustrative evidence was obtained about position and size of domains on a polypeptide chain in the case of an artificial multifunctional protein, *lac* repressor: β-galactosidase (122).

Comparison of the respective proteins from more or less related organisms may be helpful. Here one may find examples where the different functions are located on separate polypeptide chains whose molecular weights add up to about the molecular weight of the multifunctional chain. Distinct autonomous regions on a polypeptide chain usually show differing pH optima and temperature sensitivity with respect to their functions. It is possible to block one function selectively with competitive inhibitors or with amino acid specific agents without influencing the other function. Evidence along this line may come from enzyme kinetic studies or ligand binding studies. The most elegant and convincing method is to isolate and to characterize fragments that have retained their own function unimpaired. The N-terminal domains may be produced and isolated using polar mutants. Limited proteolysis is a more promising technique since in many systems the peptide bridges between domains are especially sensitive to proteolytic attack (136). Illustrative examples are DNA polymerase I from

Table 1.1 List of Multifunctional Proteins

System (source)	Structure[a] and Molecular Weights of the Subunits (daltons · 10^3)	Fused Functions	Size of the Domains (daltons · 10^3)	Cleavage in Active Fragments Possible (+/−)	References
Fused catalytic functions					
Two catalytic functions fused (bifunctional enzymes)					
Aspartokinase I: homoserine dehydrogenase I (*Escherichia coli*)	α_4 α: 86	A: aspartokinase B: homoserine dehydrogenase	A: 48 B: 55	A: + B: +	27, 29
Aspartokinase II: homoserine dehydrogenase II (*E. coli*)	α_2 α: 88	A: aspartokinase B: homoserine dehydrogenase	A: 35 B: 37	A: + B: +	30, 31
Anthranilate synthase (*Euglena gracilis*)	α α: 80	A: anthranilate synthase B: amidotransferase		−	32
Anthranilate synthase (*E. coli*, *Salmonella typhimurium*)	$\alpha_2\beta_2$ α: 62 β: 62	A: amidotransferase B: phosphoribosyl-transferase	A: 20 B: 40	A: + B: −	33, 35
Indoleglycerol-phosphate synthase (*E. coli*)	α α: 48	A: phosphoribosyl-anthranilate isomerase B: indoleglycerol-phosphate synthase	A: 25 B: 25	A: (+) B: (+)	36, 37

Tryptophan synthase (*Neurospora crassa*)	α_2 α: 75	A: cleavage of indole-glycerol-phosphate B: combination of indole and serine	A: 29 B: 45	–	38
Chorismate mutase: prephenate dehydrogenase (*E. coli*, *S. typhimurium*, *Aerobacter aerogenes*)	α_2 α: 40	A: chorismate mutase B: prephenate dehydrogenase		A: (+) B: –	39–46
Chorismate mutase: prephenate dehydratase (*E. coli*, *S. typhimurium*, *A. aerogenes*)	α_2 α: 40	A: chorismate mutase B: prephenate dehydratase		A: (+) B: (+)	41, 42, 47–49
3-Deoxy-D-arabinoheptulosonate-7-P synthase: chorismate mutase (*Bacillus subtilis*)	α_4 α: 39	A: 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase B: chorismate mutase	A: 29	A: + B: (+)	50–53
Formiminotransferase[b] cyclodeaminase (pig liver)	α_8 α: 64	A: formiminotransferase B: cyclodeaminase		A: – B: –	54, 55
Debranching enzyme (rabbit muscle)	α α: 160	A: amylo-1,6-glucosidase B: 4-α-glucanotransferase		A: – B: –	56–59
Flavocytochrome b_2 (*Saccharomyces cerevisiae*, *Hansenula anomala*)	α_4 α: 58,5	A: L-lactate-cytochrome c-oxidoreductase B: cytochrome b_2		–	60–63

Table 1.1 (*Continued*)

System (source)	Structure[a] and Molecular Weights of the Subunits (daltons · 10^3)	Fused Functions	Size of the Domains (daltons · 10^3)	Cleavage in Active Fragments Possible (+/−)	References
DNA polymerase[b] (bacteriophage T4)	α α: 115	A: DNA polymerase B: 3′-5′-exonuclease		−	64
More than two catalytic functions fused					
DNA polymerase I (*E. coli*)	α α: 109	A: DNA polymerase B: 5′-3′-exonuclease C: 3′-5′-exonuclease	A: 70 B: 35	A: + B: +	65–68
Anthranilate synthase (*N. crassa*)	$\alpha_2\beta_2$ α: 94 β: 70	A: glutaminase B: PRA isomerase[c] C: IGP synthase	A: 30 B + C: 80	A: + B + C: + B: − C: −	69, 70
Formyl-methenyl-methylenetetrahydro-folate synthetase (combined) (mammalian, *S. cerevisiae*)	α_2 α: 104	A: formyltetrahydro-folate synthetase B: methenyltetra-hydrofolate dehydrogenase C: methenyltetra-hydrofolate cyclohydrolase			71–73

Carbamylphosphate synthetase (ascites hepatoma cells)	α_4 α: 210	A: carbamylphosphate synthetase B: aspartate trans-carbamylase C: dihydroorotase	–	74
Arom-complex (*N. crassa*)	α_2 α: 150	A: DHQ synthetase[c] B: DHQase C: DHS reductase D: shikimate kinase E: EPSP synthetase	(+)	75–78
Gramicidin S-synthetase (*Bacillus brevis*)	α α: 280	specific joining of 5 different amino acids, cyclization reaction	–	79–81
Fatty acid synthetase (*S. cerevisiae*)	$\alpha_6\beta_6$ α: 185 β: 180	A: acetyl trans-acylase B: malonyl trans-acylase C. β-ketoacyl synthase D: β-ketoacyl reductase E: β-hydroxyacyl dehydrase F: enoyl reductase G: palmityl deacylase α: reactions B, E, F, G β: C, D and acyl-carrier protein	–	8, 9, 14, 82–84