SUPRAMOLECULAR STRUCTURE AND FUNCTION

DUBROVNIK, YUGOSLAVIA
SEPT 16-28 1987

Greta Pifat-Mrzljak

SUPRAMOLECULAR STRUCTURE AND FUNCTION

DUBROVNIK, YUGOSLAVIA
SEPT 16-28 1987

Editor
Greta Pifat-Mrzljak
Ruder Bošković Inst.

World Scientific
Singapore • New Jersey • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.
P. O. Box 128, Farrer Road, Singapore 9128

U. S. A. office: World Scientific Publishing Co., Inc.
687 Hartwell Street, Teaneck NJ 07666, USA

SUPRAMOLECULAR STRUCTURE AND FUNCTION

ISBN 9971-50-656-4

Printed in Singapore by JBW Printers & Binders Pte. Ltd.

Preface

In the rapidly expending field of biosciences, an enormous amount of new knowledge of the molecular basis of various biological phenomena, has emerged. The principles and methods of biophysics now provide an underpinning of all of the basic biosciences and are the rational language for discussion between scientists of different disciplines.

The International Summer Schools on Biophysics under the title **Supramolecular Structure and Fuction** held in Dubrovnik follow the successfully established interdisciplinary pattern of the schools under the same title held in 1981 and 1984 under the sponsorship of IUPAB and UNESCO.

The International School on Biophysics **Supramolecular Structure and Function** held in Dubrovnik in September 1987 had provided comprehensive discussions on a large number of subjects both for younger scientists at the doctoral or postdoctoral level interested in the molecular nature of fundamental biological entities, and for experienced scientists wishing to gain a broader insight into molecular structures and functions.

The topics discussed at the School are inter- and intramolecular interactions in biological systems, structure, organization, and function of biological macromolecules and supramolecular assemblies. A number of topics are centered around either a biological problem or a physical technique, sometimes giving an unbalanced view of the field under discussion. Some of the topics require previous knowledge of basic biophysical principles, which were then applied to gain greater insight into the molecular functions of diverse supramolecular systems.

1987 School on biophysics was designed to cover some powerful methods useful in the biological context for studing the complex biological structures and equally useful certain physical concepts for studing the interaction of drugs with biological systems. In this respect this book presents exemplary contributions to the structure of biological systems using X-ray analysis, NMR and other techniques, to thermodynamics of biomembranes etc. as the introduction to the discussion on drug interactions.

Although same of the lectures could not be prepared for publication in this volume, I hope that it contains valuable up-to-date information on various aspects of the molecular basis of life.

I wish to express my gratitude to the authors of this volume, who have contributed to the reader's understanding of biophysical problems, and perhaps provided stimulation for further studies.

The School was organized by the Yugoslav Biophysical Society, the Croatian Biophysical Society and Ruđer Bošković Institute in cooperation with UNESCO, sponsored by the International Union for Pure and Applied Biophysics (IUPAB), together with the Scientific Councils of Croatia and Yugoslavia. Financial aid to the School by these bodies is gratefully acknowledged.

Zagreb, April 1988

Greta Pifat-Mrzljak

Contents

X-RAY STRUCTURE ANALYSIS OF SYMMETRIC PROTEIN ASSEMBLIES BY DENSITY AVERAGING IN DIRECT SPACE

Including a worked example: The three-dimensional structure of the icosahedral β_{60} capsid of heavy riboflavin synthase from Bacillus subtilis.

Rudolf Ladenstein* and Adelbert Bacher**

*Max-Planck-Institut für Biochemie, D-8033 Martinsried

**Institut für Organische Chemie und Biochemie der Technischen Universität München, D-8046 Garching

I. INTRODUCTION

1. Scope of the Work

The construction of symmetric structures from asymmetric building blocks represents an important feature of nature and is studied by several disciplines of science from different viewpoints. In the field of molecular biology the symmetries of complex macromolecules are of special interest. They constitute the basis for structural organization and biological function in many cases. Maximum stability in oligomeric macromolecules is usually achieved by arranging the subunits in a symmetrical manner such that all of the subunits can form equivalent contacts.

During the past decades of research on biological macromolecules evidence has accummulated that icosahedral symmetry is an important feature which governs the self-organization of protein monomers in the formation of highly symmetric oligomeric complexes. The crystallographic work on virus structures is presently revealing the beauty, complexity and functionality of large macromolecular assemblies[1,2]. The structure analytic study on heavy riboflavin synthase described in this paper will show that icosahedral symmetry may also be of importance for the structural organization of a bifunctional enzyme complex.

The known symmetries of a complex protein oligomer may very much benefit the determination of its three-dimensional structure. The well-known Patterson search methods[3] represent efficient correlation procedures which enable the crystallographer to extract the symmetry relations among the subunits of a crystalline oligomeric macromolecule from the crystallographic intensity data alone without prior conditions. The knowledge of these symmmetries in turn allows one to use the geometric redundancy in the intensity data set of a crystalline macromolecule in order to derive new structural information by averaging[20] of electron density maps in real space.

The applicability of these methods has profitted from the development of efficient computers with high storage capacity which have made it possible to treat the vast experimental data in short time and with high accuracy.

The complex structures of highly symmetric protein molecules are fascinating in that they provide a picture of the immense potential for self-organisation inherently present in matter. Furthermore, the detailed knowledge of the structure of a macromolecular system serves as an important basis for the deeper understanding of its function. In most cases it will allow the intelligent planning of investigations with complementary methods, which may provide, together with the structure data, an insight into the complicated structure-function relations of a macromolecular assembly.

2. Heavy Riboflavin Synthase and Related Macromolecules

Heavy riboflavin synthase (HRS) from Bacillus subtilis is a bifunctional enzyme complex with a molecular weight of 10^6 Daltons. It is composed of 60 identical β subunits (M_β = 16200) which form an icosahedral capsid that encloses a trimer of α subunits (M_α = 23500)[4]. It has been shown by immunochemical methods that the immunological determinants of the α_3 trimers are not accessible for specific antibodies in the native complex $\alpha_3\beta_{60}$[5,6]. On the basis of electron microscopic data[5] and X-ray small angle scattering[7] a particle diameter of approximately 150 Å has been derived. The complex $\alpha_3\beta_{60}$ is stable only in a rather narrow pH region around pH 7. Dependent on pH and the concentration of specific substrate- and product analogous ligands (see Fig. 2) deaggregation of the native complex but also reaggregation to stable β_{60} aggregates (26 S), which are characterized by a hollow sphere shape, can occur [6]. In the absence of the stabilizing ligands polydisperse mixtures of large β aggregates are formed. The dominating species is characterized by an approximate particle diameter of 290 Å. Its architecture presumably follows the construction principles of truncated icosahedrons[6]. The

well characterized reactions leading to related β subunit assemblies are shown in Fig.1

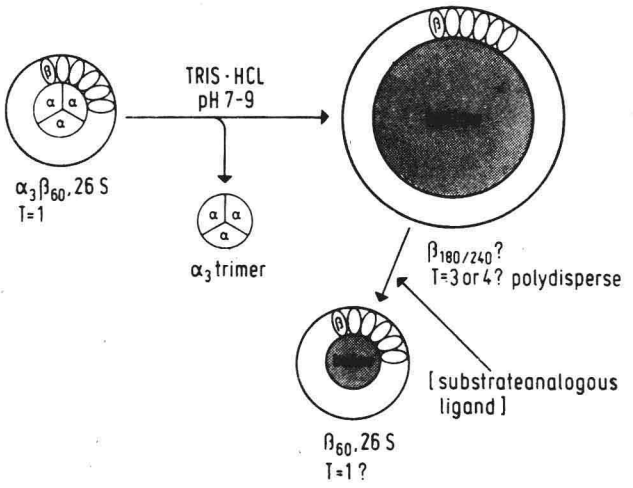

Figure 1: Deaggregation and reaggregation of heavy riboflavin synthase ($\alpha_3\beta_{60}$); formation of hollow β_{60} (26S) particles;

Thus the complex $\alpha_3\beta_{60}$ represents an ideal system for the study of protein- protein and protein-ligand interactions and the self-assembly of macromolecular systems with icosahedral symmetry.

3. The Catalytic Reaction

The bifunctional complex $\alpha_3\beta_{60}$ catalyzes the final reactions in the biosynthesis of riboflavin (vitamin B_2). Briefly, the β subunits catalyze the condensation of a 1-deoxy-tetrulose-phosphate (1) with 5-amino-6-ribitylamino-2,4 (1H, 3H) - pyrimidinedione (2) yielding 6,7-dimethyl-8-ribityllumazine (3) [8,9]. The subsequent dismutation of 3 is catalyzed by the α subunits yielding riboflavin (4) and the pyrimidinedione 2 which can be reutilized by the β subunits [4]; (Fig. 2).

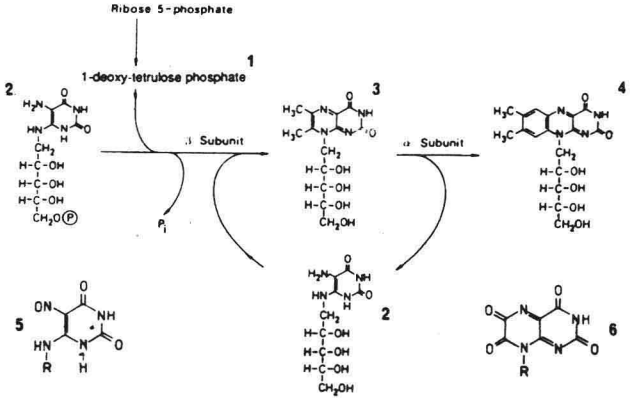

Figure 2: Biosynthesis of riboflavin;

1 = 1-deoxy-tetrulose-phosphate;

2 = 5-amino-6-(D-ribitylamino)-2,4(1H,3H)-pyrimidinedione;

3 = 6,7-dimethyl-8-(-ribityl)-lumazine;

4 = riboflavin.

Inhibitors of heavy riboflavin synthase;

5 = 5-nitroso-6-ribitylamino-2,4(1H,3H)-pyrimidinedione;

6 = 6,7-dioxo-8-ribityl-5,6,7,8-tetrahydrolumazine.

The kinetics of the catalytic steps are incompletely understood.

4. Crystals of Heavy Riboflavin Synthase

The complex $\alpha_3\beta_{60}$ could be crystallized from 1.35 M phosphate buffer pH 8.7 in the presence of 0.5 mM ligand **5** [10]. The increased pH stability under the influence of the substrate analogous ligand turned out to be crucial for successful crystallization. The crystals diffract X-rays to a resolution of 3.3 Å and belong to the hexagonal space group $P6_322$.

The unit cell dimensions are $a = b = 156.4$ Å, $c = 298.5$ Å, $\alpha = \beta = 90^0$, $\gamma = 120^0$.

As a consequence of space goup symmetry, particle dimensions and threefold particle symmetry the particle centers must sit on points with symmetry [32]. Thus the crystalline packing may be described either by hexagonal densest packing or by packing in hexagonal layers[10]. In Fig. 3 these two possibilities are shown. By electron microscopic investigation of freeze-etched 3D-crystals the packing in hexagonal layers could be verified[7].

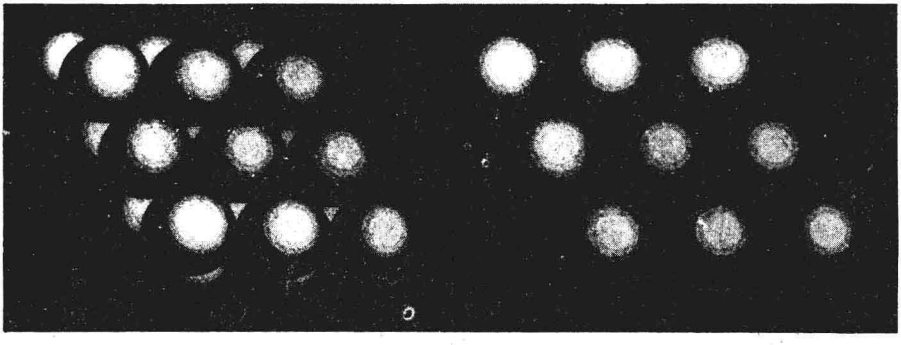

Figure 3: Sphere models showing possible crystalline packing of $\alpha_3\beta_{60}$ particles in the hexagonal unit cell; a) densest hexagonal packing; b) layer packing;

II. CRYSTALLOGRAPHIC FUNDAMENTALS

1. Diffraction of X-Rays by a Periodic Object

When a plane wave is scattered by an object, the scattered radiation may be described by the equation

$$\mathbf{F}(\mathbf{h}) = \int \rho(\mathbf{x}) \cdot \exp 2\pi i \mathbf{h}\mathbf{x} \, d\mathbf{x} \qquad (1)$$

where $\mathbf{F}$ is a complex number which represents the amplitude and phase of the scattered radiation in a direction determined by the vector

h, $\rho(\mathbf{x})$ is the scattering function at a position **x** in the object, and the integral is taken over the volume of the 3D object. For X-ray diffraction, $\rho(\mathbf{x})$ is the electron density at position **x** (excluding the effects of anomalous scattering). Thus the scattered radiation is described by the Fourier transform of the object as seen by the incident radiation. By taking the inverse Fourier transform, we get

$$\rho(\mathbf{x}) = \int \mathbf{F}(\mathbf{h}) \cdot \exp -2\Pi i \mathbf{h}\mathbf{x}\, d\mathbf{h} \qquad (2)$$

The integral is taken over the volume V* of the space spanned by the vector **h**.

We are interested in the special case in which the scattering object is a 3D crystal. The fundamental property of a crystal is that $\rho(\mathbf{x})$ is periodic in all three dimensions of space. It is known that the Fourier transform of a periodic function is zero, except when **h** is an integer multiple of the periodicity. Thus the structure factors $\mathbf{F}(\mathbf{h})$ are zero except on a three-dimensional lattice, the so-called reciprocal lattice.

The natural coordinate system for a crystal is

$$\mathbf{x} = x \cdot \mathbf{a} + y \cdot \mathbf{b} + z \cdot \mathbf{c} \qquad (3)$$

where **a,b,c** represent the basis vectors of the unit cell of the crystal. These vectors are not necessarily orthogonal but they define a three-dimensional space, which is called <u>direct space</u>. The dimension of direct space is length^3 (L^3).

In turn we may define the space of the structure factor lattice, which is related to the recorded diffraction pattern, by

$$\mathbf{h} = h \cdot \mathbf{a}^* + k \cdot \mathbf{b}^* + l \cdot \mathbf{c}^* \qquad (4)$$

with the lengths of the basis vectors $\mathbf{a}^*, \mathbf{b}^*, \mathbf{c}^*$ inversely proportional to the lengths of the basis vectors **a,b,c** of the unit cell. The space in which the structure factors are defined is generally called <u>reciprocal space</u>; its dimension is length^{-3} (L^{-3}).

In evaluation of diffraction experiments it is often necessary to transform a function defined in reciprocal space into direct space and vice versa. As we have seen these operations can be performed by Fourier transformation ($\mathcal{F}$) and inverse Fourier transformation ($\mathcal{F}^{-1}$) as

$$\mathcal{F}[\rho(\mathbf{x})] = \mathbf{F}(\mathbf{h}) \qquad (5)$$
$$\mathcal{F}^{-1}[\mathbf{F}(\mathbf{h})] = \rho(\mathbf{x}).$$

2. The Electron Density Function

If we choose $\mathbf{a}^*$, $\mathbf{b}^*$ and $\mathbf{c}^*$ to obey the Laue relations

$$\begin{aligned} \mathbf{a}\cdot\mathbf{a}^* &= 1 \quad \mathbf{a}\cdot\mathbf{b}^* = 0 \quad \mathbf{a}\cdot\mathbf{c}^* = 0 \\ \mathbf{b}\cdot\mathbf{a}^* &= 0 \quad \mathbf{b}\cdot\mathbf{b}^* = 1 \quad \mathbf{c}\cdot\mathbf{c}^* = 0 \\ \mathbf{c}\cdot\mathbf{a}^* &= 0 \quad \mathbf{c}\cdot\mathbf{b}^* = 0 \quad \mathbf{c}\cdot\mathbf{c}^* = 1 \end{aligned} \qquad (6)$$

the vector products in the integrals (1) and (2) simplify to $\mathbf{h}\cdot\mathbf{x} = hx + ky + lz$. With these definitions we are able to normalize (1) and (2) to reflect the contents of one unit cell; we get

$$\mathbf{F}(\mathbf{h}) = V \iiint \rho(\mathbf{x})\cdot\exp[2\Pi i(hx + ky + lz)]dxdydz \qquad (7)$$
$$\rho(\mathbf{x}) = 1/V \iiint \mathbf{F}(\mathbf{h})\cdot\exp[-2\Pi i(hx + ky + lz)]dhdkdl \qquad (8)$$

The discrete nature of $\mathbf{F}(\mathbf{h})$ allows the conversion from an integral to a sum in (8)

$$\rho(\mathbf{x}) = 1/V \sum_h\sum_k\sum_l \mathbf{F}(\mathbf{h})\cdot\exp[-2\Pi i(hx + ky + lz)] \qquad (9)$$

Equation (9) represents the well-known electron density equation which can be calculated by inverse Fourier transformation of the scattered waves, described by the structure factors $\mathbf{F}(\mathbf{h})$. The structure factor $\mathbf{F}(\mathbf{h})$ is characterized by an amplitude $|\mathbf{F}(\mathbf{h})|$ and a

phase $\alpha(\mathbf{h})$ according to

$$\mathbf{F}(\mathbf{h}) = |\mathbf{F}(\mathbf{h})| \cdot \exp i\alpha(\mathbf{h}) \qquad (10)$$

The phase information $\alpha(\mathbf{h})$ is lost in the diffraction experiment; phases are generally determined by methods such as single (SIR) and multiple (MIR) isomorphous replacement and phase extension (see chapter III.6) The amplitudes $|\mathbf{F}(\mathbf{h})| = \text{const.} \cdot \sqrt{I(\mathbf{h})}$ are obtained from measurement of the crystallographic intensities in diffraction patterns (e.g., photographic rotation method[11], area detectors[12] and diffractometers[13]. The set of all symmetry independent $I(\mathbf{h})$ represents the unique intensity data set of a crystal. Crystallographic intensity data of native and derivative crystals of heavy riboflavin synthase are shown in Table I

Table I Intensity data statistics ($\mathbf{F}^2 > 1.0\ \sigma$)

Derivative	Resolution	Measurements	Independ. reflections	measured/possible reflections	R_{merge}[%]
ITAN	∞–3.3	75100	27700	0.850 to 3.3Å	13.1
AuCN	∞–3.2	87250	32200	0.673 to 3.2Å	12.8
CMAA	∞–3.6	71300	24800	0.767 to 3.6Å	13.9
WAC	∞–3.4	38200	23000	0.741 to 3.4Å	12.8
WP	∞–3.6	19000	13700	0.434 to 3.6Å	9.4
LUMO	∞–3.6	32500	18300	0.684 to 3.6Å	11.0

ITAN, native crystals; $R_{merge} = \Sigma\Sigma|\langle I_h\rangle - I_{hi}| / \Sigma N_h \langle I_h\rangle$, where $\langle I_h\rangle$ is the average intensity of N_h measurements and I_{hi} is the individual intensity of a reflection h;

LUMO, functional derivative obtained by soaking the native crystals with 1mM of the dioxolumazine (Ligand **6**) in 1.75 M potassium phosphate buffer, pH = 8.7, at 20° C.

3. Formal Description of Symmetries

By definition the crystallographic symmetries represent the set of symmetry elements valid in a crystal; they relate the asymmetric units of the crystal cell. The noncrystallographic or local symmetry elements are confined to the asymmetric unit and can be described by the set of symmetry operations which are valid in the asymmetric unit. In the case of an oligomeric protein the asymmetric unit of a crystal cell may contain more than one copy of a subunit. The positions of these subunits in direct space are defined by the symmetry operations of the assymetric unit.
Generally we may define a symmetry operation in 3D space by way of a linear transformation including a 3x3 Matrix **R** and a translation vector **t** as

$$\mathbf{x}' = \mathbf{R}\cdot\mathbf{x} + \mathbf{t} \qquad (12)$$

and

$$\rho(\mathbf{x}') = \rho(\mathbf{x}) \qquad (13)$$

for all positions **x** within a crystal or its asymmetric unit.

4. The Symmetry of Icosahedral Polyhedrons

The ancient Greek mathematicians already knew that only five regular polyhedrons can exist, the so-called platonic solids. These are tetrahedron, octahedron, icosahedron, cube and dodecahedron. Nowadays it is well established that geometry and symmetry of these bodies represent an important principle which governs the self-assembly of protein subunits in quite a large number of cases. A regular icosahedron (Fig. 4) is constructed from 20 equilateral triangles and possesses 6 fivefold ($n = 72^{\circ}$), 10 threefold ($n = 120^{\circ}$) and 15 twofold ($n = 180^{\circ}$) rotation symmetry axes. All of these axes intersect at a common point which represents the center (origin) of the particle.