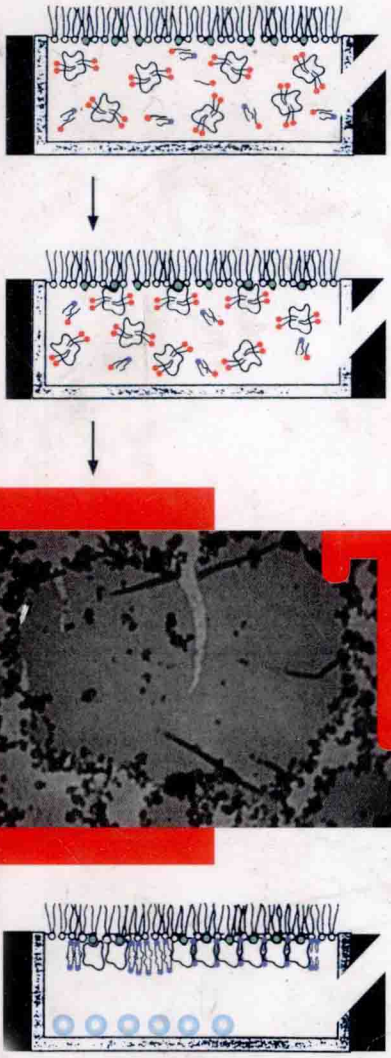

# Strategies for Two-dimensional Crystallization of Proteins Using Lipid Monolayers

Jens Dietrich

Vénien-Bryan

Imperial College Press

# Strategies for Two-dimensional Crystallization of Proteins Using Lipid Monolayers

Jens Dietrich
Catherine Vénien-Bryan
University of Oxford, UK

ICP
Imperial College Press

*Published by*

Imperial College Press
57 Shelton Street
Covent Garden
London WC2H 9HE

*Distributed by*

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
*USA office:* 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
*UK office:* 57 Shelton Street, Covent Garden, London WC2H 9HE

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

**STRATEGIES FOR TWO-DIMENSIONAL CRYSTALLIZATION OF PROTEINS USING LIPID MONOLAYERS**



ISBN 1-86094-428-0

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore by World Scientific Printers (S) Pte Ltd

# Strategies for Two-dimensional Crystallization of Proteins Using Lipid Monolayers

*"... mare ... omne oleo tranquillari, et ob id urinantes ore spargere, quoniam mitiget naturam asperam lucemque deportet."*

Caius Plinii Secundi
*Naturalis historia*, Liber II CVI.
Mirabilia fontium et fluminum

*"All sea water is made smooth by oil, and so divers sprinkle oil from their mouth because it calms the rough element and carries light down with them."*

Pliny the Elder
Roman Naturalist AD 23–79,
*Natural History*, Book 2 Chapter CVI

## Acknowledgments

We would like to thank the following people:

- Christine Ziegler and Peter Bond for help with some figures;
- Elizabeth Hewat, Kevin Leonard, Richard K. Bryan and Michael Troke for critically reading the manuscript.
- The Wellcome Trust and Human Frontier Science Programme for financial support

Jens Dietrich
Catherine Vénien-Bryan
July 2004

# Introduction

After sequencing the human genome and the genome of many other organisms, the age of "structural genomics" has started. It is now possible to take the genetic sequence and infer the amino acid sequence of all proteins from it. Unfortunately this does not necessarily tell us something about the function of these proteins. Even for the organism that is genetically best understood, *Escherischia coli*, the function of ca. 40% of its protein is not known. Under the structural genomics initiative, many groups set out to determine as many protein folds as possible. A protein fold is a stretch of amino acids that fold into a defined secondary structure motif. These structural prototypes adopt similar structures, but do not necessarily have detectable sequence homology. At the moment, about 1,000 different folds are known and it is estimated that a few thousand different folds exist. Often X-ray crystallography of 3D protein crystals and nuclear magnetic resonance (NMR) studies of concentrated protein solutions are the only source of structural information at atomic level, and most of the work in structural genomics is limited to water-soluble proteins. Membrane proteins and larger protein complexes are mostly excluded from these efforts because they are less amenable to this "high-throughput" approach.

Membrane proteins constitute 25% of all proteins, therefore it is crucial to improve our understandings of these proteins. The "Protein Data Base" (PDB) stores ca. 28,000 structures of water-soluble proteins (about 18,000 of these are unrelated), but only 87 structures of membrane proteins are available (of which 51 are unrelated) and only 93 structures originate from proteins with a molecular weight >250kDa.

The handling of membrane proteins and large protein complexes is much more demanding than for water-soluble proteins. The reasons for this are:

- Membrane proteins are usually difficult to express in large quantity for structural analysis;
- Monodispersity and stability of the purified protein or large protein complexes are often difficult to control;
- The localization of membrane proteins in the lipid bilayer requires that they exhibit hydrophilic and hydrophobic surfaces.

Therefore solubilization and purification of membrane proteins necessitates the use of detergent for masking the hydrophobic area. As a consequence of this dual property — hydrophobic and hydrophilic — the total hydrophilic surfaces available to provide a good crystal contact necessary for 3D crystallization are very limited. The hydrophobic domain, which is masked by detergent micelles, does not play a major role in the crystal contact. The growth of 3D crystal of membrane proteins is therefore a complicated task. In this context, electron microscopy of single particles (for protein with a molecular weight >250 kDa) and 2D crystals is a powerful technique for which, in some cases, no other alternative approach is possible. Thanks to advances in electron microscopy instrumentation, specimen preparation and image processing, this technology is beginning to satisfy the demand for structures and allow learning about mechanisms at atomic level. The atomic structures of light harvesting complex, bacteriorhodopsin, and tubulin have now been solved, crucial elements of secondary structure have been revealed in several membrane proteins (aquaporin, rhodopsin, gap junctions and $Ca^{2+}$ and $H^{+}$-ATPase) and a novel viral fold of the hepatitis B core protein has been determined through the application of this technique.

Nevertheless the resolution obtained by this technique is often in the medium range and the ability to combine structures of macromolecular complexes derived by electron microscopy with X-ray or NMR structures of their components allows the reconstruction of molecular machines and large multi-protein complexes in considerable detail. Two-dimensional crystals have been most successfully used to obtain high and medium resolution structural information by electron microscopy.

In this book we offer an overview of the technique of crystallization of proteins on lipid monolayers. The structural information is then obtained by imaging the 2D crystals using electron microscopy and image processing. This method allows any soluble or membrane protein from very small molecular weight to large complexes such as viruses to be crystallized in two dimensions. Strategies to adsorb, concentrate, orient and organize proteins or macromolecules on supports suitable for electron microscopic observation and with fluidity properties similar to biological membranes will be presented and discussed. Biophysical techniques to monitor and improve the process of crystallization will be detailed.

# Contents

# Chapter 1

# Two-dimensional Crystallization on Lipid Monolayers

## 1.1. Overview

In 1971 it was first shown by Fromherz (1971) that an ordered arrangement of protein can be generated underneath a lipid monolayer. Ferritin molecules were observed in a regular arrangement after adsorption to a lipid film. Limited lateral mobility and heterogeneity in the lipid monolayer resulted in only moderately ordered protein in these early experiments. About 10 years later, for the first time Kornberg's group was capable of obtaining 2D crystals of an antibody using a monolayer of lipid hapten (Uzgiris and Kornberg, 1983). Since then, this 2D crystallization technique has been successfully employed for a variety of different proteins (see Appendix). These are mainly water-soluble proteins, but recent studies have shown that this technique is also applicable to membrane proteins. This may provide a much needed extension in the repertoire of membrane protein crystallization, thus improving our structural knowledge of this protein class.

The crystallization on lipid layers is an elegant method because it is possible to work with very dilute protein solutions and still generate a locally high concentration of protein constrained in 2D. Nonetheless, the proteins retain sufficient mobility to allow for organization into crystalline 2D arrays by lateral diffusion. Lipid monolayers can be spread (driven by surface tension) over the whole air/water interface of a drop to form a flat, one molecule thick film. This provides a substrate for protein binding, leading to a layer of closely packed proteins at the interface which can be organized into a 2D crystal suitable for structure determination by electron crystallography.

The first step in the crystallization process is the adsorption of the protein to a lipid monolayer (Fig. 1.1). This limits the protein to a few orientations relative to the lipid plane which facilitates crystallization. The hydrophilic headgroup of the lipid is responsible for this function and operates as a recognition element for the protein in one of two ways. The first involves electrostatic interactions of a charged lipid layer with the protein allowing non-specific binding to a surface layer. The second occurs by specific binding of protein to a surface monolayer. This is achieved by interaction of the protein with a ligand attached to the polar headgroup of the lipid. Both types of interactions lead to a densely packed protein layer at the lipid/water interface. It is possible to obtain a locally high concentration of protein in the order of 500–1,000mg/mL from a very dilute solution (10–100μg/mL) (Kornberg and Darst, 1991).

Lipids provide a substrate for protein binding and therefore a functional basis of the crystallization technique. In general, lipids consist firstly of a hydrophilic headgroup which can carry a charge or a functionalized ligand group, for example a $Ni^{2+}$-NTA group, that can be utilized for protein binding via a His-tag. The second part of a lipid is a long hydrophobic tail which usually consists of two acyl chains. It is necessary for the

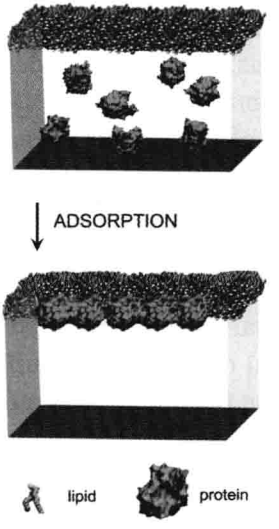


**Fig. 1.1** The method of surface crystallization for soluble proteins.

molecule to accommodate its hydrophobic and hydrophilic part in a suitable environment. Therefore, amphiphilic molecules like lipids have a tendency to self-organize. In solution, they form lipid bilayers in the form of vesicles or tubes; at the air/water interface they form lipid monolayers. The hydrophobic tails are sheltered from the water by burying within the bilayer or by pointing into the apolar air. In this way interaction of each part with the opposite phase is minimized. This is favorable because the acyl chains are not able to form hydrogen bonds with water. The organization of the monolayer is driven by the hydrophobic effect associated with the lipid tail groups of the lipid molecules which also determines the stability and fluidity of the monolayer. The structure of this water insoluble moiety is responsible for the physical properties of the lipid layer, because the long hydrophobic chains interact extensively with each other. The length and saturation of the acyl chain as well as branches in the chain influence the fluidity of the lipid at a given temperature. Fluidity of a lipid monolayer is given by intramolecular contacts between the hydrophobic lipid chains. This interaction is dependent on the shape of the hydrophobic tails of the lipids. A *cis*-double-bond introduces a kink in the acyl chain which sterically hinders packing of the hydrocarbon chains in the same fashion as a branched chain and both result in an increased fluidity of the lipid (see Sec. 3.1). It is clear from the literature that the lipid monolayer being in the fluid state is the most favorable condition for the growth of 2D crystals. Although closely packed protein molecules have been found on monolayers in the solid state, 2D crystals were not observed under these conditions (Darst *et al.*, 1991a; Mosser and Brisson, 1991).

The physical properties of the monolayer system are determined mainly by the chemical composition of the lipids, the temperature and the composition of the underlying buffer. In order to allow protein crystallization these parameters have to permit the lateral diffusion of the protein molecules attached to the monolayer. To achieve a favorable physical state of the lipid layer it has often proven useful or even essential to use mixtures of different lipids. As these additional lipids usually do not carry a functional group, they are generally referred to as diluting lipids. A reason for the dependence on diluting lipids might be the difference in surface covered by proteins and the much smaller lipids. A phospholipid occupies about 50 times less area than a 100kDa globular protein.

Therefore, one protein molecule can interact with many lipid molecules. Thus the composition of the lipid layer is of high importance for crystallization, especially in the case of non-specific interactions, as it determines the local electrostatic milieu. Furthermore, it is possible to adjust the fluidity properties for a given monolayer by mixing functionalized lipids with diluting lipids of different structure.

Application of the surface crystallization technique for the creation of 2D crystals always bears the possibility of producing hexagonally closely packed protein, which might be mistaken for protein crystals at low resolution. Some proteins crystals like cholera toxin show a true hexagonal space group (Kornberg and Ribi, 1987), other denser areas of the same protein on the grid were found to exhibit a hexagonal diffraction pattern (Ludwig *et al.*, 1986). Further investigation indicated that the diffraction pattern was due to the presence of hexagonally closely packed protein molecules, and not due to a crystal. Surface crystallization trials generally lead to a closely packed layer of protein and are therefore prone to produce this kind of artifact, which rarely lead to 2D crystals. Figure 1.2 shows different organizations of protein which are common for surface crystallization trials.

The observation of growing 2D crystals on lipid monolayers can also give insight into the fundamental processes during crystallization (Ku *et al.*, 1993). Due to reduced dimensionality, aspects of the transition between disordered and ordered states can be studied more easily. 2D crystals have also been used to promote epitaxial growth of 3D crystals from 2D crystals from lipid layers. In this case, 2D crystals serve as nuclei for the formation of 3D crystals. Furthermore, it has been established that the lipid layers alone can trigger epitaxial crystal growth (Hemming *et al.*, 1995).

In conclusion, there are some general requirements for 2D crystallization of proteins:

- limiting the protein diffusion to a plane;
- a high concentration of the protein in the plane;
- orientation of the protein;
- providing mobility of the protein within the plane to allow sampling of various interaction arrangements.

The physical properties of the lipids spread at the air/water interface, the protein adsorption and binding capacity at the interface, the

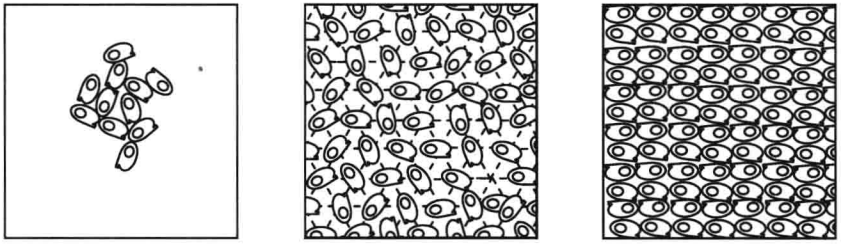


**Fig. 1.2** Different stages of protein organization on lipid films: (A) aggregates without order, (B) closely packed hexagonal arrays, (C) 2D crystal.

lipid reorganization induced by protein binding and the crystallization process of the proteins have been extensively studied using numerous biophysical techniques. Some of these techniques include film balance measurements, ellipsometry, light scattering microscopy, epifluorescence microscopy, fluorescence spectroscopy, fluorescence microscopy, phase contrast microscopy, Brewster angle microscopy (BAM), electron microscopy (EM), atomic force microscopy (AFM). More recently, other methods have been developed such as X-ray reflectivity, X-ray grazing incidence diffraction, neutron reflectivity, shear modulus measurements and scanning near field optical microscopy (SNOM). Some of these techniques will be presented in chapter 3.

## 1.2. Non-specific Adsorption through Electrostatic Interactions

Lipids can contain headgroups with positive, negative, or neutral charges (Fig. 1.3). The attraction of opposite electrical charges provides the basis for electrostatic interactions. Charged lipids can be used to create a charged surface that can interact with the surface potential of a protein. This type of interaction can be compared to processes involved in ion-exchange chromatography:

> Proteins carry positive and negative charges, according to the acidic or basic side chains of single amino acids. At acidic pH values the side chain lysine, arginine and histidine are

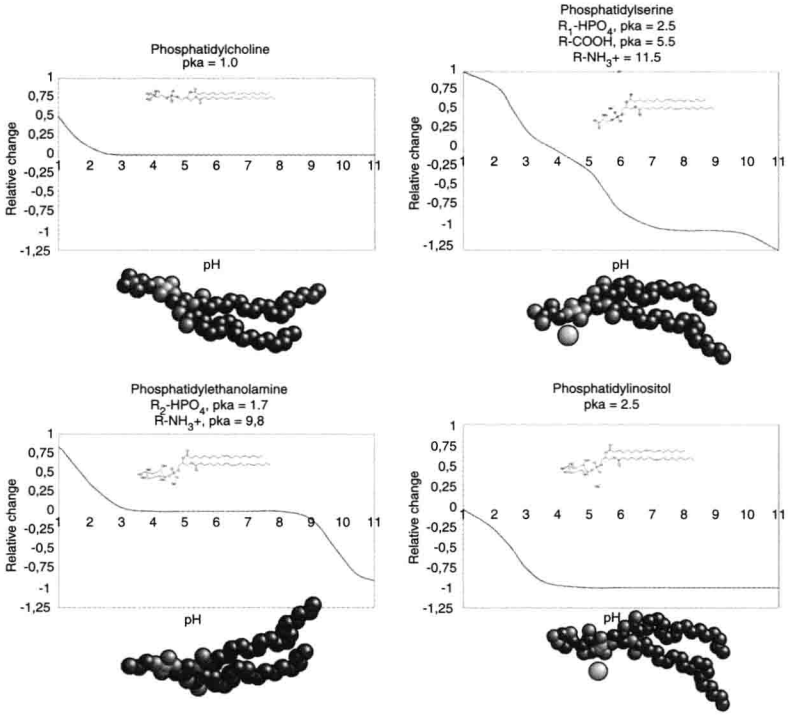


**Fig. 1.3** Chemical structure, surface representation and relative charge as a function of pH of some common phospholipids.

protonated and a protein shows a cationic behavior. In contrast, at basic pH the negative charges of aspartic acid and glutamic acid render the protein anionic. Therefore, the net charge of a protein is dependent on the pH of the surrounding buffer and the number of exposed charged amino acids on the surface of the protein. The overall charge of a protein is described by the pI value. If the pH of the buffer is at the pI, the net charge of the protein is zero and therefore the capacity of the protein for electrostatic interactions is low, unless there is an unequal distribution of charged amino acids on the surface of the protein, which leads to regions with a positive or negative charge.

This seems to be the case for α-actinin which forms 2D crystals on a positively charged lipid layer at the isoelectric point of the protein