PART 3. TECHNIQUES OF MOLECULAR BIOLOGY

Chapter 3.1. Preparing recombinant proteins for X-ray crystallography

S. H. Hughes and A. M. Stock

3.1.1. Introduction

Preparing protein crystals appropriate for X-ray diffraction usually requires a considerable amount of highly purified protein. When crystallographic methods were first developed, the practitioners of the art were compelled to study proteins that could be easily obtained in large quantities in relatively pure form; the first proteins whose structures were solved by crystallographic methods were myoglobin and haemoglobin. Unfortunately, some of the most interesting proteins are normally present in relatively small amounts, which, while it did not prevent crystallographers from dreaming about their structures, prevented any serious attempts at crystallization. Recombinant DNA techniques changed the rules: it is now possible to instruct a variety of cells and organisms to make large amounts of almost any protein chosen by the investigator. Not only can specific proteins be expressed in large quantities, recombinant proteins can be modified in ways that make the task of the crystallographer simpler and can, in some cases, dramatically improve the quality of the resulting crystals. It is not our intention in writing this chapter to provide either a methods manual for those interested in expressing a particular protein or a complete compendium of the available literature. The literature is vast and complex, and, as we will discuss, the problems associated with expressing a particular protein are often idiosyncratic, making it difficult to provide a simple, comprehensive, methodological guide. What we intend is to discuss issues (and problems) relevant to choosing methods appropriate for preparing recombinant proteins for X-ray crystallography. In this way, we hope to help readers understand both the extant problems and the available solutions, so that, armed with a general understanding of the issues, they can more easily confront a variety of specific projects.

Fortunately, there are a large number of additional resources available to those who are interested in expressing and purifying recombinant proteins, but lack the expertise. These include numerous methods books (*e.g.* on molecular biology: Sambrook *et al.*, 1989; Ausubel *et al.*, 1995; on protein purification: Abelson & Simon, 1990; Scopes, 1994; Bollag *et al.*, 1996), useful reviews of the literature (cited throughout), formal courses (such as those offered by Cold Spring Harbor Laboratory), meetings (*i.e.* IBC's International Conference on Expression Technologies, Washington DC, 1997) and a specialized journal (*Protein Expression and Purification*). The pace of methodological development is rapid, and company catalogues, publications and web pages can provide extensive, useful, up-to-date information. In many cases, a convenient source of information is a nearby researcher whose own research depends on expressing and purifying recombinant proteins. Those who are serious about preparing recombinant proteins for crystallography, but have little or no experience, are strongly urged to avail themselves of these resources. In many cases the help of a knowledgeable colleague is the most valuable resource. In general, the literature provides a much better guide to what will work than what will fail; quite often, in designing a good strategy to produce a recombinant protein that is suitable for crystallography, it is more important to understand the potential pitfalls. Discussion with an experienced colleague is usually the best way to avoid the most obvious errors.

Section 3.1.2 gives an overview of the problem, Section 3.1.3 discusses engineering an expression construct, Section 3.1.4 discusses expression systems, Section 3.1.5 discusses protein purification and Section 3.1.6 discusses the characterization of the purified product.

3.1.2. Overview

The idea that underlies the problem of expressing large amounts of a recombinant protein is straightforward: prepare a DNA segment that, when introduced into an appropriate host, will cause the abundant expression of the relevant protein. However, as the saying goes, 'The devil is in the details.' Not only is it necessary to design the appropriate DNA segment, but also to introduce it into an appropriate host such that the host retains and faithfully replicates the DNA. The DNA segment must contain all of the elements necessary for high-level RNA expression; moreover, the RNA, when expressed, must be recognized by the translational machinery of the host. The recombinant protein, once expressed, needs to be properly folded either by the host or, if not properly folded in the host, by the experimentalist. If the protein is subject to post-translational modifications (cleavage, glycosylation, phosphorylation *etc.*) and the experimentalist wishes to retain these modifications, the appropriate signals must be present and the chosen host must also be capable of recognizing the signals. Once the recombinant protein is expressed, assuming it is reasonably stable in the chosen host, the protein must be purified; as we will discuss, recombinant proteins can be modified to simplify purification. Once purified, the quality of the protein preparation must be evaluated to ensure it is both relatively homogeneous and monodisperse.

While this chapter will be limited to discussions of the basic strategies for creating an expression vector, expressing the protein and purifying and characterizing the product, molecular biological methods can be used in other ways that are relevant to crystallography. In some cases, a protein in its natural form is not suitable for crystallization. Crystallographers have long used proteolytic digestion and/or glycolytic digestion to produce proteins suitable for crystallization from ones that are not. Such techniques have been used to good effect on recombinant proteins; however, the ability to modify the segment encoding the protein makes it possible to alter the protein in a variety of ways beyond simple enzymatic digestions. Specific examples of such applications are described in Chapter 4.3.

Unfortunately, no single strategy for producing proteins for crystallization appears to be universally successful. Any parti-

references