Protein research, proteomics and applied enzymology
Editorial overview
Richard R Burgess and Bernard Witholt

The outpouring of powerful new advances in protein technology and commercial enzyme application continues. This issue reviews some of the recent developments in these areas.

There are now three major thrusts to modern protein research. First is the detailed analysis of protein structure and function at the basic research level and often at the single molecule level, with the aim of understanding proteins as molecular machines. Second is the emergence of proteomics — the effort to understand the structure, function, interaction partners, processing variants, and regulation of all of the proteins encoded by an entire genome. The most vigorous participants in this area are those involved in drug discovery. Third is the manipulation and improvement of enzymes and optimization of their efficient use as commercial biocatalysts for carrying out chemical conversions.

High-throughput screening, the ability to robotically screen huge libraries of potential enzyme or cell growth inhibitors for lead compounds for pharmaceutical development, has spawned ‘high-throughput thinking’ and dramatically broadened the scope and expectations of studies of protein–protein interaction, binding specificity, and quantitative determination. In addition, high-throughput protein expression, purification, crystallization trials, and structure determination are being actively pursued by many pharmaceutical and genomics/proteomics companies in search of new structures and drug targets for pharmaceutical discovery. Simultaneous protein profiling of thousands of human proteins from normal and diseased biological samples offers tremendous future promise for diagnosis, tailored treatment, and treatment monitoring of millions of patients.

To prepare large quantities of recombinant proteins for research or industrial use, good expression systems are needed. Cereghino, Cereghino, Ilgen, and Gregg (pp 329–332) review recent advances in the production of recombinant proteins in fermenter cultures of the methylotrophic yeast Pichia pastoris. This yeast can easily be manipulated genetically; it can express proteins at high levels both intracellularly and extracellularly, and it can perform many ‘higher eukaryotic’ protein modifications such as glycosylation, disulfide bond formation, and proteolytic processing. This host is certainly a viable choice and has produced functional proteins in cases where expression has not been successful in the bacteria Escherichia coli, the yeast Saccharomyces cerevisiae or baculovirus-infected insect cells.

Immunoaffinity chromatography is one of the most powerful fractionation steps available for protein purification; however, it is often difficult to elute bound protein without using harsh or denaturing elution conditions. Burgess and Thompson (pp 304–308) describe the development of methods to identify monoclonal antibodies that bind antigens tightly, but release these under gentle, non-denaturing conditions. One example is the polyol-responsive antibody, which enables the immunoaffinity purification of labile, multi-subunit enzyme complexes with high yield and high biological specific activity. Gentle immunoaffinity chromatography has implications for emerging proteomic
applications; allowing identification of new protein–protein interaction partners and isolation of protein complexes more amenable to crystallization and structure determination. This article also reviews the growing use of purification tags, especially those based on antibody epitopes.

Protein microarrays usually contain specific monoclonal antibodies, designed for proteomic studies to simultaneously measure the level of many proteins or analytes in a biological sample. Petach and Gold (pp 309–314) describe a potentially very powerful approach that does not use monoclonal antibodies, but instead uses photoaptamers. An aptamer is a short nucleic acid with a sequence selected for tight binding to a given protein or analyte. A photoaptamer is a special aptamer that, upon binding to the analyte, positions a photoactive nucleoside in the sequence adjacent to a tryptophan residue in the analyte such that, upon UV irradiation, a highly specific covalent crosslink is formed. This allows more stringent washing, decreasing the background of non-specific analyte binding to the array. This review discusses the implications of a high-density panel of such analyte specific reagents and considers what types of arrays will yield the most useful biological information in medical proteomics applications.

Arrays of peptides of defined sequence facilitate studies of molecular recognition events and identification of biologically active peptides. The article by Reimer, Reineke and Schneider-Mergener (pp 315–320) reviews applications of peptide arrays in epitope mapping, characterization of protein–protein interactions in general, and enzyme–substrate or enzyme–inhibitor interactions, both with macroarrays and microarrays.

Hofmann and Muir (pp 297–303) review recent advances in the application of expressed protein ligation (EPL) to protein engineering. Recombinantly expressed proteins are linked to polypeptides containing biophysical probes, post-translational modifications, or unnatural amino acids. One important application is the ligation of a heavy atom isotope labeled small segment of a larger protein to the unlabeled remainder. This approach allows NMR structural determination of the labeled segment in the context of the native full-length protein, which may be too large for NMR studies if the whole protein were labeled.

Changes in the conformation of a protein during substrate binding or formation of protein–protein binding pairs are usually followed spectroscopically. Heyduk (pp 292–296) describes the use of fluorescence resonance energy transfer (FRET) and luminescence resonance energy transfer (LRET) to measure distances between protein domains on nanosecond timescales. These techniques allow measurements on single protein molecules, in vivo as well as in vitro. New ways of introducing fluorescence probes into proteins, newly developed fluorescence probes, and improvements in fluorescence signal detection greatly expand the range of application of FRET.

Proteomics strives to analyze the complete protein complement of a given organelle, tissue, or cell. The review by Patton, Schulenberg and Steinberg (pp 321–328) compares two techniques presently being used for this purpose: two-dimensional gel electrophoresis (2DGE) and isotopic-coded affinity tagging (ICAT). 2DGE separates many hundreds of proteins in biological samples based on isoelectric point and size, and determines approximate amounts of each protein resolved. However, it is difficult to resolve very complex samples, achieve consistent lab-to-lab reproducibility, and analyze very high or low molecular weight or very alkaline or hydrophobic proteins. ICAT is a relatively new method that involves derivatizing a protein mix with a mass-labeled, cysteine-selective reagent. The precise mass of proteins or peptides in the mixture by mass spectrometry through comparison with an internal standard protein mix derivatized identically but with a different mass-labeled reagent. This ingenious method allows quantification of proteins in a biological sample relative to a standard sample and gets around the basic limitation of mass spectrometry: that peak height is not proportional to the amount of different proteins in a mixture. However, the review points out that ICAT is blind to proteins lacking cysteine residues and presents data showing that neither ICAT nor 2DGE provide comprehensive coverage on a proteome-wide scale.

Enzymes have traditionally been used in the food, feed, and consumer industries. van Beilen and Li (pp 338–344) present an overview of enzyme technology indicating that rapid technological developments are also stimulating the chemistry and pharma industries to embrace enzyme technology; a trend strengthened by concerns regarding health, energy, raw materials and the environment. In developing new applications, all steps in the process sequence must be considered, from biocatalyst screening to product purification. It is interesting, however, that, of these various operational steps, the careful formulation and utilization of the biocatalyst remains a major determinant of success.

One of the fascinating discoveries of the last few decades has been the demonstration that enzymes can function in nearly dry organic solvents. This is important because many organic chemical substrates and products are sparingly soluble or insoluble in aqueous solvents. Lee and Dordick (pp 376–384) describe the many potential advantages of biocatalysis in organic solvents, and discuss the present limitations of this system, especially with respect to attainable reaction rates. Surprisingly, enzyme activity can often be dramatically improved through lyophilization of aqueous biocatalyst solutions in the presence of organic or inorganic excipient molecules, such as non-buffer salts, macrocycles (e.g. cyclodextrins), and certain denaturants and surfactants. This knowledge has led to a broad array of techniques for generating active, stable, enantio-selectively and regio-selectively tailored enzymes for synthetic transformations in non-aqueous solvents. A related review by Marhuenda-Egea and Bonete (pp 385–389) discusses the use of enzymes from extreme halophiles in organic solvents. The authors focus on the use of enzymes trapped inside reverse micelles to maintain activity under unfavorable conditions, including low salt concentrations.
Even when an enzyme catalyzes a desired reaction, the formulation of a functional biocatalyst particle is a crucial component in developing an effective bioprocess. van Roon, Beeftink, Schroen, and Tramper (pp 398–405) discuss advances in assessing the intra-particle biocatalyst distribution as a tool in rational formulation. The development of specific immunolabeling techniques, high-quality image digitizing devices, and enhanced computing power are making the assessment and optimization of the internal biocatalyst profile an integral part of biocatalytic particle optimization.

Novozymes is the largest producer of industrial enzymes worldwide. Ole Kirk and co-workers (pp 345–351) describe the effects of directed evolution and gene shuffling in altering and improving enzyme properties in expanding the range of industrial enzyme applications. Industrial enzyme applications continue to be based on hydrolytic enzymes, with proteases and carbohydrateases being the most important groups. These are now being modified and optimized to function effectively (and inexpensively) in detergents at lower temperatures.

Lipases constitute a most important group of biocatalysts for biotechnological applications, including the synthesis of biopolymers and biodiesel, the production of enantiopure pharmaceuticals, agrochemicals and flavor compounds. Jaeger and Eggert (pp 390–397) summarize the expression and folding of recombinant lipases, optimization of activity by directed evolution, and several specific biotechnological applications of lipases.

Improved enzyme stability is often highly desirable, whether to facilitate purification and storage, to increase the ease of data collection during X-ray crystallography, or to improve performance of a commercial enzyme. There are many ways to select mutations that increase the stability of proteins, including rational design, functional screening of randomly generated mutant libraries, and comparison of naturally occurring homologous proteins. Such thermostability engineering is discussed in detail by van den Burg and Eijsink (pp 333–337), who conclude that there is no single general solution to the engineering of thermostability in mesostable proteins. Early strategies, such as converting monomeric proteins to presumably more stable multimers, have not led to the expected successes. In contrast, charged surface residues that form networks of electrostatic interactions have turned out to be surprisingly important in protein stabilization. The authors point out that although the protein engineer’s toolbox is expanding and successful examples of the engineering of increased protein stability abound, selection is complicated by the lack of knowledge of the process of thermal inactivation and the diverse strategies employed by proteins to achieve stability.

The development of biocatalytic industrial processes has traditionally depended on the adaptation of these processes to the typical properties of enzymes (fragile and difficult to work with). Today, there is more emphasis on the modification of enzymes to suit the industrial reactor conditions, as discussed by Huisman and Gray (pp 352–358), who also stress the integration of biocatalysis and chemical synthesis to develop optimized processes. Hence, in their view, enzyme modifications are process-guided, which they illustrate with the development of a process for the improved production of compactin, a precursor of pravastatin, one of the present top blockbuster drugs.

The impacts of the use of industrial enzymes in the large-scale production of useful chemicals in Europe and Japan are reviewed by Schmid, Hollmann, Park, and Bühler (pp 359–366) and by Ogawa and Shimizu (pp 367–375), respectively. Many European chemical industries are looking towards the life sciences for new products and processes. In addition to biocatalytic steps in production of organics and agricultural chemicals, applications are moving into drug synthesis, plastic materials and synthetics and in the near future into products like soaps, cleaners, inorganics, adhesives, and specialty chemicals.

Japan has a long tradition of industrial fermentations for the large-scale production of amino acids and coenzymes. These early successes have, in the past few decades, led to the evolution of enzymatic processes for industrial chemicals. Ogawa and Shimizu (pp 367–375) describe some of these new processes, including hybrid processes based on combinations of enzymatic and chemical steps. Thus, the merits of biocatalysis and chemical synthesis can be compared directly, sometimes resulting in substitutions of chemical with biocatalytic steps, as was the case for the synthesis of D-β-hydroxyphenylglycine, produced since 1995 at a rate of 2,000 tons per year. A dozen similar biocatalytic processes, all introduced in the past decade, are described. Enzymes discussed include those for production of amino acids, nucleotides and chiral compounds and enzymes for the conversion of nitriles to commodity chemicals, and for transformation of fatty acids into polyunsaturated and conjugated fatty acids.

Finally, several of the reviews emphasize the continued importance of empirical approaches to complement molecular design efforts. As Huisman and Gray (pp 352–358) suggest, and the recent work on enzymes in solvents described by Lee and Dordick (pp 376–384) amply illustrates, we are sometimes blinded by erroneous but well-established assumptions. Hence, efforts at design, though highly educational, often fail. Difficulties in developing clear rules for the modifications needed to stabilize proteins imply that empirical approaches will remain important in producing improved enzymes, as discussed by van den Burg and Eijsink (pp 333–337). Similarly, in improving industrial enzymes, no single approach will be dominant. Instead, as Kirk and co-workers (Novozymes; pp 345–351), Ogawa and Shimizu (pp 367–375), and Huisman and Gray (Maxygen; pp 352–358) have indicated, a combination of design and random modifications followed by screening will continue to be needed to develop enzymes adapted to industrial processes.