50th Anniversary Celebration Collection

Special Section on Perspective on Drug Metabolism and Disposition, Part I—Minireview

Recombinant Technologies Facilitate Drug Metabolism, Pharmacokinetics, and General Biomedical Research

Joseph M. Cronin and Ai-Ming Yu

Department of Biochemistry and Molecular Medicine, UC Davis School of Medicine, Sacramento, CA (J.M.C., A.-M.Y.)

Received June 23, 2022; accepted March 9, 2023

ABSTRACT

The development of safe and effective medications requires a profound understanding of their pharmacokinetic (PK) and pharmacodynamic properties. PK studies have been built through investigation of enzymes and transporters that drive drug absorption, distribution, metabolism, and excretion (ADME). Like many other disciplines, the study of ADME gene products and their functions has been revolutionized through the invention and widespread adoption of recombinant DNA technologies. Recombinant DNA technologies use expression vectors such as plasmids to achieve heterologous expression of a desired transgene in a specified host organism. This has enabled the purification of recombinant ADME gene products for functional and structural characterization, allowing investigators to elucidate their roles in drug metabolism and disposition. This strategy has also been used to offer recombinant or bioengineered RNA (BioRNA) agents to investigate the posttranscriptional regulation of ADME genes. Conventional research with small noncoding RNAs such as microRNAs (miRNAs) and small interfering RNAs has been dependent on synthetic RNA analogs that are known to carry a range of chemical modifications expected to improve stability and PK properties. Indeed, a novel transfer RNA fused pre-miRNA carrier-based bioengineering platform technology has been established to offer consistent and high-yield production of unparalleled BioRNA molecules from Escherichia coli fermentation. These BioRNAs are produced and processed inside living cells to better recapitulate the properties of natural RNAs, representing superior research tools to investigate regulatory mechanisms behind ADME.

SIGNIFICANCE STATEMENT

This review article summarizes recombinant DNA technologies that have been an incredible boon in the study of drug metabolism and PK, providing investigators with powerful tools to express nearly any ADME gene products for functional and structural studies. It further overviews novel recombinant RNA technologies and discusses the utilities of bioengineered RNA agents for the investigation of ADME gene regulation and general biomedical research.

Introduction

The understanding of drug metabolism and pharmacokinetic (DMPK) processes represents a critical and necessary step in the development of new drugs. Elucidating these processes and understanding how they differ between species is key in translating basic scientific research into an efficacious and safe clinical treatment. The importance of DMPK in drug development and safety was first made widely evident by the tragedy of thalidomide, a morning sickness drug developed with rat models that caused severe birth defects in humans (Speirs, 1962; Boylen et al., 1963). While the cytochrome P450 (CYP or P450) enzymes were identified and characterized spectroscopically in 1962, widespread appreciations of P450s in xenobiotic metabolism and interspecies differences in drug disposition had not yet been established (Omura and Sato, 1962). Further research with thalidomide determined that its teratogenic effect could be attributed to toxic metabolites discovered in the urine of treated animals (Gordon et al., 1981). Notably, this toxic bioactivation of thalidomide was absent in rat liver microsomes and only observed in rabbits, primates, and humans (Gordon et al., 1981). This tragedy and subsequent discovery illustrated the importance of interspecies differences in drug metabolism and resulting toxicity, driving forward DMPK research and influencing key regulatory law still in effect today.

This work was supported by the National Institutes of Health National Institute of General Medical Sciences [R35-GM140835] and National Cancer Institute [R01-CA252330 and R01-CA225968]. J.M.C. was supported by a National Institutes of General Medical Sciences-funded Pharmacology Training Program grant [T32-GM144303].

No author has an actual or perceived conflict of interest with the contents of this article.

dx.doi.org/10.1124/dmd.122.001008.

ABBREVIATIONS: ADME, absorption, distribution, metabolism, and excretion; BioRNA, bioengineered RNA; cDNA, complementary DNA; CRISPR, clustered regularly interspaced short palindromic repeats; CYP or P450, cytochrome P450 monooxygenase; DDI, drug-drug interaction; DMPK, drug metabolism and pharmacokinetics; FDA, Food and Drug Administration; mRNA, messenger RNA; miRNA, microRNA; ncRNA, noncoding RNA; NRF2, nuclear factor erythroid 2-related factor 2; ori, origin of replication; PK, pharmacokinetic; rDNA, recombinant DNA; siRNA, small interfering RNA; SRP, signal recognition particle; tRNA, transfer RNA; VDR, vitamin D receptor.
Soon afterward, key advancements in molecular biology and the subsequent birth of recombinant DNA (rDNA) technologies (Fig. 1) would equip DMPK researchers with the tools to investigate unprecedented research questions in drug metabolism and disposition (Guengerich, 2020). This was not limited to CYP enzymes but also included other genes and proteins involved in absorption, distribution, metabolism, and excretion (ADME) processes. Nucleotide sequencing technologies enabled the generation of complementary DNA (cDNA) libraries for CYP enzymes (Negishi et al., 1981; Gonzalez and Gelboin, 1992) (Fig. 1) and other ADME gene products for use in molecular cloning (Maniatis et al., 1983; Fothergill-Gilmore, 1993; Guengerich, 2020; Patil et al., 2022). Generated cDNA sequences were used to produce recombinant P450 enzymes through heterologous expression in model systems such as yeast (Saccharomyces cerevisiae) and Escherichia coli (E. coli) (White et al., 1984; Oeda et al., 1985; Unger et al., 1986; Gonzalez et al., 1988a,b; Umeno et al., 1988; Podust et al., 2001; Yu et al., 2002; Williams et al., 2003; Schoch et al., 2004). With previous research limited to the use of liver microsomes or P450s purified from the tissues (Lu and Coon, 1968; Estabrook, 2003; Coon, 2005), the development of rDNA strategies propelled the field forward by enabling the characterization of individual P450 enzymes in vitro (Fig. 1) (Oeda et al., 1985; Gonzalez et al., 1988a,b; Umeno et al., 1988; Barnes et al., 1991; Larson et al., 1991b; Li and Chiang, 1991; Scott et al., 2001; Guengerich, 2020). This approach proved to be a significant boon to the fields of DMPK, empowering researchers to interrogate key differences between human P450 reactions and those relevant to preclinical animal models (Guengerich, 2020). rDNA technologies have proven invaluable in the human P450 reactions and those relevant to preclinical animal models DMPK, empowering researchers to interrogate key differences between individuals, populations, and species can have drastic variations as well as broader racial differences such as those in in observed in poor CYP2D6 and CYP2C19 metabolizers (Nakamura et al., 1985; Gonzalez et al., 1988b). Further complicating matters, CYP enzyme expression can be induced through exposure to a wide variety of drugs and environmental chemicals (Graham and Lake, 2008). Although a variety of induction mechanisms have been elucidated, this largely occurs through xenobiotic interaction with nuclear receptors to drive an increase in CYP gene transcription (Dickins, 2004). Natural products have also been identified to inhibit critical drug-metabolizing CYP enzymes rather than induce them, with the most widely known example being grapefruit juice’s inhibition of CYP3A4 (Guo and Yamazoe, 2004). Similar to expression level and catalytic activity, induction of CYP genes in response to environmental chemicals or drugs can vary widely between species as well as between individuals (Tang et al., 2005; Graham and Lake, 2008). Understanding these differences in ADME gene expression, protein activity, and gene regulation in response to chemicals is critical to designing safe drugs that account for species and individual differences in drug metabolism and toxicity.

Although significant advancements have been made, continuing to broaden our understanding of ADME gene regulation and function in DMPK is necessary to inform the development of safe and efficacious medications. When analyzed by exome sequencing, ADME genes were found to have significantly more variability and possible variants compared with non-ADME genes (Hovelson et al., 2017). ADME gene variations between individuals, populations, and species can have drastic effects on DMPK, therefore affecting therapeutic safety and efficacy. These variations between species can have a dramatic impact on the success of drug development, decreasing the face validity of preclinical animal models and negatively affecting whether scientific research can be translated into an effective clinical therapy. Approximately 54% of novel therapeutics failed in clinical development from 1998 to 2015, with the large majority of failures attributed to inadequate efficacy or safety (Hwang et al., 2016). In addition to this, therapies that were successfully approved took an average time of 17 years for research discoveries to reach the clinic (Morris et al., 2011). Recent updates to US Food and Drug Administration (FDA) guidelines have also removed the requirement that all drugs be tested in animals, instead welcoming more representative models of human disease such as organoid or organ-on-a-chip approaches (Wadman, 2023). Keeping these shortcomings in
mind, it is clear that novel technologies and approaches are needed to further elucidate ADME gene regulation and function in DMPK. Advancement of these fields is critical to educate researchers and improve the success rate of developing drugs, as well as shorten the lag time observed bringing a drug from “bench to bedside.” This mini-review examines rDNA technologies in the context of DMPK research, highlighting historical successes expressing recombinant proteins for structural and functional studies as well as discussing new recombinant RNA approaches that have offered unique RNA molecules to investigate ADME gene regulation and novel research questions.

rDNA Technology

rDNA technology is characterized by altering genetic material outside an organism with the goal of transforming desired characteristics into a living organism or its products. Recombinant technologies would not be imaginable without those that laid the scientific foundation of how we understand heredity, genetics, and molecular biology. In this way, all modern successes of rDNA technologies stand on the shoulders of giants like Gregor Mendel that first began to elucidate these key genetic processes (Mendel, 1865). These researchers broadened the scientific understanding of genetics and DNA, enabling the development of critical technologies that set the stage for genomic sequencing, molecular cloning, and modern rDNA technologies.

rDNA Technology: A Brief History. In 1952, Joshua Lederberg proposed the term “plasmid” for an extrachromosomal genetic material that had been recently observed (Fig. 1) (Lederberg, 1952). At the time, no one could predict that the plasmid would one day become the workhorse at the center of countless recombinant technologies across a diverse range of industries. We know now that plasmids are circular, double-stranded DNA molecules naturally present in the cytoplasm of bacteria and other prokaryotes that are capable of replicating independently from the host cell (Fothergill-Gilmore, 1993; del Solar et al., 1998). Plasmids encode the information needed for their own replication and can confer specific characteristics to the host organism such as antibiotic resistance (Fothergill-Gilmore, 1993). However, research at the time of their discovery was hampered by a lack of viable strategies to introduce these circular double-stranded DNA molecules into cells (Patil et al., 2022). This research problem would lead to a monumental breakthrough for rDNA technologies by driving collaboration between experts in plasmid technology and targeted DNA manipulation (Patil et al., 2022). This collaboration and the intersection of these critical techniques resulted in the construction of the first biologically functional recombinant plasmid in vitro (Cohen et al., 1973) (Fig. 1). Restriction endonucleases EcoRI and EcoRII were used to cut the double-stranded plasmid, leaving single-stranded 3’ and 5’ sticky ends that could be joined together by DNA ligase to generate a circular double-stranded DNA molecule and functional plasmid (Cohen et al., 1973). This key experiment combined separate E. coli plasmid species pSC101 and pSC102, which contain genes encoding for resistance to tetracycline and kanamycin, respectively. The resulting biologically functional plasmid conferred resistance to both antibiotics, which was used to identify positive transformants via antibiotic treatment (Cohen et al., 1973). This represents the first use of what is commonly known now as a selection marker gene: a critical characteristic of plasmids and other cloning vectors that allows for selection of positively transformed cells. This approach utilizing selective advantage to identify positive transformants has proven invaluable to rDNA technologies and is still widely used today.

The following year, in 1974, this innovative strategy was successfully used to achieve the replication and transcription of eukaryotic Xenopus laevis DNA in E. coli (Morrow et al., 1974). Detection of X. laevis rDNA in E. coli transformants from this experiment demonstrated that E. coli could be used to clone genes with eukaryotic species of origin (Morrow et al., 1974). The next logical step was to test whether this approach could be employed to not only transcribe a eukaryotic gene but also to express and produce human proteins. Shortly after, the human hormone somatostatin was successfully cloned in E. coli using a pBR322 plasmid vector, demonstrating that rDNA technologies could be used to produce functional human proteins (Itakura et al., 1977) (Fig. 1). Only a few years later, Genentech’s recombinant human insulin named Humulin gained US FDA approval and entered the US market (Goeddel et al., 1979). This landmark advancement in rDNA technology would serve as the inspiration for the development of countless biotechnologies, showcasing vector-driven transformation in microorganisms like E. coli as a robust strategy to achieve the expression of a human gene product. With 621 biologic drugs currently FDA approved in the United States, it is clear these initial advances in rDNA technologies have inspired decades of innovative scientific research (FDA, 2021).

Expression Vectors: Characteristics and Examples. Since a bacterial plasmid was first used to express a eukaryotic gene in E. coli, our understanding of plasmids and other cloning vectors has greatly advanced. The bacterial plasmid was the first vector identified and optimized for the purpose of expressing recombinant gene products. However, there are also components present in the genomes of bacteriophages, yeasts, and viruses that can be used to accomplish effective transfer, cloning, and expression of a transgene in a heterologous system (Vieira and Messing, 1982; Patil et al., 2022). Specialized vectors that are optimized for transgene expression in a desired host organism are referred to as expression vectors and often contain additional non-coding sequences such as promoters, enhancers, and exons from the parent sequence that improve the transcription or expression of the gene product (Patil et al., 2022). Although vectors vary in their source, size, structure, capacity, and replication efficiency, all of them must retain several key attributes. To facilitate the discussion of key vector principles and their impact on successful transformation and expression of a gene insert, we refer to the widely used and iterated-upon plasmid pBR322 as an example (Fig. 2). This was one of the first vectors constructed with the intent of generating and selecting for recombinant clones, with pBR322 and its derivatives such as the pUC plasmids being commercially available and still widely used in rDNA technologies today (Balázs et al., 1986).

Notable sequences used to construct the pBR322 plasmid include the ampicillin resistance gene (ampR), a promoter from the CoIE1-related plasmid named pRSF214, and the tetracycline resistance gene (tetR) from the pSC101 plasmid discussed earlier (Cohen et al., 1973; Bolivar et al., 1977a,b) (Fig. 2). As established, cloning and expression vectors must contain a selection marker gene that confers a selective advantage to vector-bearing host cells (Cohen et al., 1973; Patil et al., 2022). Selection markers commonly encode antibacterial resistance or metabolic advantages, providing a survival advantage to positive transformants so they can be selected for through antibiotic treatment or nutrient deprivation (Bolivar et al., 1977b; Fothergill-Gilmore, 1993; Patil et al., 2022) (Fig. 2). Strategies to select for recombinant clones are not only limited to those providing a survival advantage to positive transformants, with other successful selection approaches utilizing targeted DNA insertion to functionally disrupt a gene (Bolivar et al., 1977b). This can confer traits to positive transformants that allow for their negative selection, such as tetracycline sensitivity in tetR disruption and the formation of colorless colonies when the gene encoding the alpha protein of beta-galactosidase is disrupted (Bolivar et al., 1977b; Vieira and Messing, 1982).

In addition to containing a selection marker gene, a suitable expression vector must also be able to replicate in the host organism of choice.
Fig. 2. Recombinant protein expression vectors: Characteristics and examples in P450 research. (A) First-generation vector pBR322 contains an ampicillin resistance gene (AmpR) and a tetracycline resistance gene (TcR), both shown in green. These genes serve as selection markers to identify positive transformants upon antibiotic treatment. The pBR322 plasmid contains a ColE1 ori site, which is illustrated in yellow. This ori site is recognized by E. coli host machinery to enable replication of the plasmid in this model organism. The repressor of primer (rop) gene illustrated in red encodes the rop protein, which limits the copy number of this vector. (B) An optimized expression vector derived from pCWOri+ is used to construct target plasmid toward heterologous P450 expression in E. coli (von Wachenfeldt et al., 1997). This vector contains two tac promoter sequences that are under transcriptional control of the lac repressor protein, which is illustrated in red (Barnes et al., 1991). The lac repressor is encoded by the lacIq gene shown in green, which suppresses transcription until an inducing agent such as isopropyl β-D-thiogalactopyranoside is added. The vector and recombinant P450 gene insert are digested with endonucleases NdeI and EcoRI, enabling insertion and ligation of the transgene downstream of the two inducible tac promoters. This pCWOri+ derived expression vector also possesses two ori sites shown in yellow, which include the previously discussed ColE1 ori as well as an F1 ori that provides a recognition site for M13 bacteriophages.
In plasmids, phagemids, and bacteriophage-based vectors, this is ensured through the inclusion of an origin of replication (ori) site (Patil et al., 2022). The ori site is recognized by host replication machinery, serving as a point of assembly for the components involved in DNA replication (del Solar et al., 1998; Patil et al., 2022). Some ori sites are specific to a certain host bacteria, while others are able to be recognized by the replication machinery of multiple species (Patil et al., 2022). The ori site in the pBR322 plasmid is derived from the ColE1 plasmid native to E. coli, making it a suitable choice for cloning and expression in this host organism (Bolivar et al., 1977b) (Fig. 2). The inclusion of more than one ori site in a cloning vector can allow replication across multiple different host species, expanding the flexibility of vectors for use in other model systems. The ori sites within expression vectors are often optimized for more efficient recruitment of replication components, with these sequences having a dramatic effect on the copy number of the vector (Friehs K, 2004; Patil et al., 2022). Another determinant of pBR322’s copy number is encoded by the repressor of primer (rop) gene shown in red (Cesareni et al., 1982; Banner et al., 1987) (Fig. 2). The protein encoded by this gene regulates plasmid copy number by promoting interaction of inhibitory RNA components that prevent primer maturation to modulate plasmid replication (Banner et al., 1987; Helmer-Citterich et al., 1988; del Solar and Espinosa, 2000). This limits the copy number of pBR322, a disadvantage in cloning that would be addressed by future derivatives of this vector.

Expression vectors must also contain multiple cloning sites that serve as unique locations for restriction endonucleases to recognize, enabling the specific digestion of the plasmid at these sites to generate a linearized double-stranded DNA molecule. The transgene of interest is digested using the same restriction enzymes to enable gene insertion into the plasmid. The compatible ends generated by restriction digest are ligated together by DNA ligase, creating a recombinant vector that can undergo replication using host machinery (Fothergill-Gilmore, 1993; del Solar et al., 1998) (Fig. 2). The pBR322 plasmid has been optimized to include a variety of multiple cloning sites for gene insertion, with many of them falling within critical segments of the vector, including both antibiotic resistance genes, the ampR promoter, and the origin of replication (Bolivar et al., 1977b) (Fig. 2). The functional disruption of the tetR gene via DNA insert that was discussed earlier utilizes the unique multiple cloning sites present within this gene. This allows for identification of host cells bearing the plasmid that contains the transgene DNA insert, with positive pBR322 transformants that successfully inserted and ligated the DNA fragment via HindIII, Sall, or BamHI endonucleases no longer having a functional tetR gene (Bolivar et al., 1977b; Patil et al., 2022). This creative utilization of multiple cloning sites to generate specialized selection markers highlights the vast utility of restriction endonucleases, not only proving invaluable to the targeted insertion of a transgene but also in the customization of expression vectors and their features. While there are many context-specific applications of these enzymes and the cloning sites they recognize, this general strategy of enzyme-mediated transgene insertion, ligation, and replication has been widely applied in the cloning and expression of countless gene products as therapeutics and for functional study, including those integral to ADME processes such as P450 enzymes (Oeda et al., 1985; Unger et al., 1986; Gonzalez et al., 1988b) (Fig. 2).

In addition to the discussed components, an effective expression vector must also promote the efficient expression of the inserted transgene in the host system. Inclusion of a strong promoter that is compatible with the host species is critical to achieve adequate expression of a recombinant gene product (Fothergill-Gilmore, 1993). The promoter sequence must also be compatible with the modality of the gene product, with different promoters being used depending on whether the recombinant gene product is a protein, messenger RNA (mRNA), or small interfering RNA (siRNA; Patil et al., 2022). Some specialized expression vectors can also be designed with dual-promoter systems that aim to further increase the strength of transgene expression (Oztürk et al., 2017). These dual-promoter systems can consist of two twin promoters activated under the same conditions, as well as two nonidentical promoters activated during separate steps to aid production processes in biotechnology (Oztürk et al., 2017). Similarly, the inclusion of a strong termination sequence following the promoter and transgene can enhance expression and represents an important component of an efficient vector (West and Proudfoot, 2009; Gasanov et al., 2015; Patil et al., 2022). As recombinant technologies have continued to advance, expression vectors have become more readily adaptable tools that can be customized to efficiently express a wide variety of transgene products across many host systems. Recently developed expression vectors may also aid the purification of recombinant proteins through the installation of a polypeptide affinity tag. Once purified, the affinity tag can be cleaved by tobacco etch virus protease to return the recombinant protein to its native conformation (Rocco et al., 2008). This strategy has been used to generate several novel expression vectors from the pET series capable of expressing and more easily purifying recombinant proteins (Rocco et al., 2008).

Like the choice of expression vector, choosing an appropriate host organism for transgene expression is a critically important step when utilizing rDNA technologies. Different model organisms possess distinct characteristics that can make them best suited to a particular application or gene product (Fothergill-Gilmore, 1993). Model organisms for recombinant expression can vary in their yield capacity, post-synthetic modification capability (glycosylation, acetylation), ability to secrete active proteins, and ability to express specific gene products or modalities. Common model species for transgene expression include bacteria (E. coli, Lactococcus lactis, Pseudomonas), yeast (Saccharomyces cerevisiae), mammalian cells (CHO, human embryonic kidney, COS), and insect cells (Sf9) (Koga et al., 1985; Oeda et al., 1985; Wurm, 2004; Chen, 2012). Research efforts continue to improve upon these expression systems and remedy their shortcomings. One of the most relevant examples of this is in E. coli, where advancements have enabled glycosylation of protein products and the use of antibiotic-free selection plasmids (Chen, 2012).

Applications of rDNA Technology in DMPK Research. Since the inception of rDNA technology, innovative researchers have been developing creative ways to adapt the rDNA platform to address their research questions with cDNA-expressed P450s (Fig. 1). One critical example of this was the development of site-specific mutagenesis methods utilizing M13 bacteriophage cloning vectors (Hutchison et al., 1978; Zoller and Smith, 1982). This method is used to elicit specific and targeted changes to a gene of interest for transformation into E. coli and other model organisms (Hutchison et al., 1978; Zoller and Smith, 1982). With the ability to introduce targeted single-nucleotide changes to any DNA fragment, researchers were able generate novel mutants for functional study. Often this approach is used to investigate the role of individual amino acid residues in contributing to the function or interactions of a specific protein. In DMPK research, this method has been used to generate novel P450 mutants to elucidate the role that specific residues play in protein behavior and enzyme function (Wilderman et al., 2012). Site-directed mutagenesis has also been widely applied in the characterization of other ADME related proteins such as ABC transporters and multidrug efflux pumps (Yao et al., 2013; Cox et al., 2018). Distinct from site-directed mutagenesis, random mutagenesis has also been used in DMPK research to generate novel CYP enzymes capable of withstanding harsh environmental conditions and performing bio-transformations previously unseen in nature (Giver et al., 1998; Joo et al., 1999; Brandenberg et al., 2017; Arnold, 2018). Through many
cycles of random mutagenesis and screening, beneficial mutations are accumulated that alter metabolic function and often promote an intrinsic promiscuity or side-reactivity of the natural enzyme (Arnold, 2018). This strategy developed by Frances Arnold is termed directed evolution, and it was awarded the 2018 Nobel Prize in chemistry. This innovative approach has been used to develop novel CYP enzymes capable of catalyzing many new and nonnatural reactions such as those valuable to medicinal chemistry or biotechnology processes (Brandenberg et al., 2017; Arnold, 2018).

Another valuable application of rDNA technologies in DMPK research has been improving the field’s access to robust research tools. A prominent example of this is the production of recombinant antibodies and nanobodies utilizing phage or yeast display libraries (Marks et al., 1991; McMahon et al., 2018). Specific and sensitive antibodies are critical reagents for any biomedical research, serving as important probes to detect and quantify target antigens. Recombinant antibodies may also serve as inhibitors to block specific protein-protein interactions, proving useful in elucidating the role of specific protein epitopes and interactions in determining protein function (Basu et al., 2019).

Previous methods of antibody production relied on hybridoma technology, requiring mature B cells to be harvested from an animal immunized with the antigen (Basu et al., 2019). Because this relies on host immunization, the antibody generated is in response to host presentation of the antigen via the major histocompatibility complex. The antigen must undergo proteolysis to a smaller peptide for host presentation to occur, and because of this the antibodies generated by hybridoma methodology do not always successfully recognize the native conformation of the antigen they target (Roche and Furuta, 2015; Basu et al., 2019). However, when generating recombinant antibodies by phage display, the antigen of interest is present in its native conformation and is less likely to undergo alteration compared with hybridoma approaches (Basu et al., 2019). This represents a distinct advantage of recombinant antibody production when compared with previous hybridoma methods, especially when considering the importance of conformation in ADME-related protein function. Additionally, recombinant antibody production is significantly cheaper, does not require the use of animals, and allows for the inclusion of additional features such as an amino acid tag to aid the purification process (Basu et al., 2019).

This success utilizing rDNA technology to develop superior research tools has not only occurred with antibodies but has also extended to other modalities such as microRNAs (miRNAs), siRNAs, and aptamers (Ho and Yu, 2016). miRNAs are understood as epigenetic regulators with a vast array of targets and functions, with one miRNA able to regulate several mRNA transcripts at the posttranscriptional level (Ali Syeda et al., 2020; Yu et al., 2020). Understanding miRNA’s role in ADME gene regulation and function is critical to addressing a comprehensive picture of DMPK. The field of miRNA research has been somewhat limited by the widespread usage of chemically synthesized miRNA mimics as research tools. These commercially available miRNA mimics carry extensive chemical modifications that are needed to improve their stability and PK properties (Khvorova and Watts, 2017; Yu et al., 2019). The way in which these synthetic mimics are modified is not always disclosed, and modifications between manufacturers can vary in identity, location, and number (Yu et al., 2019). Although the nucleotide sequence is unaffected, these inconsistent modification practices generate miRNAs that are constitutively different from natural miRNAs and may lead to conflicting results between investigators (Yu et al., 2019; Diener et al., 2022; Traber and Yu, 2023).

To address these shortcomings, the Yu laboratory has developed an innovative bioengineering platform to express and purify recombinant miRNA molecules in E. coli (Li et al., 2021; Tu et al., 2021). Echoing the success of biologic proteins, these miRNAs are synthesized and folded inside living cells and better recapitulate the structures, modifications, and properties of endogenous miRNAs and thus represent superior tools for DMPK research and as potential therapeutics (Yu et al., 2019). Later in this review, we discuss the optimization of this recombinant strategy as well as its utilization in elucidating the impact of endogenous miRNA on ADME gene regulation and DMPK.

In addition to developing improved tools and reagents for research, advancements in rDNA technologies have also enabled novel approaches to generate improved cellular and animal models for the study of drug metabolism and human disease (Gonzalez and Yu, 2006; Dorr et al., 2017; Sundhari et al., 2019; Lu et al., 2021). One prominent advancement is the 2020 Nobel Prize winner in chemistry, an adaptable gene editing platform that utilizes clustered regularly interspaced short palindromic repeats (CRISPR) associated with a Cas9 endonuclease (Westermann et al., 2021). This promising technology commonly known as CRISPR-Cas9 has gained widespread popularity over the past decade, with many ex vivo gene therapies that use the platform currently undergoing clinical trials across a variety disease states (Li et al., 2020). In the context of DMPK research, this targeted gene editing technology has been used to generate valuable in vivo and in vitro knockouts to characterize individual ADME gene and allele function (Religa et al., 2021). Examples of ADME gene products knocked out or modified for functional study using CRISPR-Cas9 include CYP enzymes, ABC transporters, nuclear receptors, and solute carrier proteins (Karlgren et al., 2018; Wei et al., 2018; Liang et al., 2019; Karakus et al., 2020; Salanga et al., 2020; Lu et al., 2021). This approach can accomplish specific knockout of individual or multiple alleles of a target gene, allowing researchers to interrogate their role in xenobiotic metabolism, drug-drug interactions (DDIs), and PK (Dorr et al., 2017; Lu et al., 2021). While the CRISPR/Cas9 platform holds great promise to enable research and treat genetic disorders, future iterations of the technology must continue to improve upon its accuracy and precision as well as address concerns surrounding its unintended epigenetic impact (Doudna, 2020).

Recombinant ADME Proteins for Functional and Structural Studies

Before rDNA technologies gained widespread usage in the development of therapeutica and research tools, these recombinant advancements were used in the functional and structural study of ADME gene products such as CYP enzymes (Fig. 1). Indeed, early research of P450 enzymes before rDNA technology was limited to approaches such as preparing microsomes and purifying individual CYPs from animal tissues (Lu and Coon, 1968). This strategy proved effective in reconstituting soluble cytochrome P450 enzymes with intact catalytic activity, allowing researchers to study P450-mediated biotransformation of endogenous compounds and xenobiotics alike (Lu and Coon, 1968; Lu et al., 1969; Guengerich, 1977). Additionally, these early works from Lu and Coon identified the necessary components that constitute catalytic activity in liver microsomes: P450 hemoprotein, cytochrome c reductase, and a heat-stable factor (Lu and Coon, 1968). These research efforts broadened the field’s understanding of in vivo drug metabolism by identifying numerous unique forms of soluble P450 enzymes, elucidating key components critical to catalytic activity, and illuminating characteristic differences between species (Lu and Coon, 1968; Haugen et al., 1975; Ryan et al., 1975; Guengerich, 1977).

Microbial P450CAM: A Window Into P450 Structure and Function. Despite some success with liver microsomes, the DMPK field was in need of a more suitable model system for the structural and
functional characterization of individual P450 enzymes (Poulos et al., 1982). This role would be filled by the microbial cytochrome P450CAM, a camphor-metabolizing monooxygenase that was identified in *Pseudomonas putida* (Gunsalus and Wagner, 1978; Poulos et al., 1982, 1987) (Table 1). This enzyme was purified for study prior to the advent of rDNA technologies and instead was isolated from bacterial culture according to procedures optimized by Gunsalus and Wagner (Gunsalus and Wagner, 1978). P450CAM was the first CYP enzyme that had been purified in both soluble and crystalline forms, making it an ideal biologic model to study P450 structure and function (Gunsalus and Wagner, 1978; Poulos et al., 1982, 1985). High-resolution crystal structures were generated of P450CAM both bound to camphor and in the absence of its substrate, providing valuable insight into the substrate binding site and the structure-function relationship of the enzyme (Poulos et al., 1986, 1987). This crystallography data also elucidated the secondary structural features of the polypeptide, illustrating important structural features of the enzyme such as the oxygen/heme binding pocket as well as a metal cation interaction that may stabilize active site formation (Poulos et al., 1987).

These advancements in P450CAM visualization were also paired with amino acid sequencing techniques, further improving the detail of structural and functional characterization and enabling the comparison of P450CAM to eukaryotic P450 sequences (Poulos et al., 1987). Compared P450 sequences were found to be fairly homologous, with residues that form the oxygen binding area being highly conserved between species (Poulos et al., 1987). Amino acid sequencing also revealed that eukaryotic P450s uniquely possess a hydrophobic NH2 terminus, a domain that was speculated at the time to contribute to the membrane solubility observed in mammalian P450s (Poulos et al., 1987). While bacterial cytochrome enzymes like P450CAM are easily solubilized in the cytoplasm, eukaryotic P450 enzymes reside in the membranes of the endoplasmic reticulum and to a lesser extent the mitochondria (Lee et al., 1981). The functional role that this hydrophobic NH2 terminus plays in mammalian P450 membrane association and its effect on catalytic activity would later be elucidated using rDNA technologies (Oeda et al., 1985; Sakaguchi et al., 1984; Sakaguchi et al., 1987). P450CAM has also served as an effective model system to study the role of accessory proteins, with crystal structures of the enzyme complexed to its redox partner putidaredoxin providing insight into the structural changes induced by putidaredoxin and how they may facilitate electron transfer (Hiruma et al., 2013; Tripathi et al., 2013). Even today, P450CAM is used as a model system to further investigate the allosteric roles of accessory proteins and how they may impact conformational dynamics (Poulos and Follmer, 2022).

Although P450CAM was purified from a gram-negative bacterium, its extensive study aided our understanding of mammalian P450 enzymes by enabling the thorough characterization of well-conserved domains and functions (White et al., 1984; Black and Coon, 1987; Poulos et al., 1987; Gonzalez et al., 1988a,b; Umeno et al., 1988; Gonzalez and Gelboin, 1989)."
1992). With isolation of pure P450 enzymes from human tissue being an
arduous and time-consuming process, this P450cam model system was
invaluable in establishing initial structure-function relationships and elic-
itating the basic catalytic mechanism conserved across nearly all P450s
(Guengerich, 2020; Poulos and Follmer, 2022). P450cam was also one of the first P450 enzymes cloned for recombinant expression, utilizing a
plasmid-derived vector under strict promoter control to produce catalyt-
ically active P450 monooxygenase in *P. putida* and *E. coli* host systems
(Koga et al., 1985; Unger et al., 1986). This was not the only microbial
P450 enzyme investigated utilizing rDNA technologies, as the crystal
structure of the prokaryotic form of 14α-demethylase (CYP51) present in
*Mycobacterium tuberculosis* was also elucidated this way (Podust et al.,
2001) (Table 1). While CYP51 exists in most species as a membrane-
bound enzyme that is localized to microsomes, the *Mycobacterium tuber-
culus* genome posesses a uniquely soluble form of the enzyme that was
expressed in *E. coli* for easier determination of the crystal structure
(Bellamine et al., 1999; Podust et al., 2001). This newfound structural
information would inform rational drug design of improved antifungal
agents that target CYP51.

cDNA Directed Expression of Mammalian P450 Proteins. Not
long after the crystal structure and amino acid sequence of P450cam
were elucidated, cDNA strategies gained widespread usage in the study of
mammalian CYP enzyme structure and function in drug metabolism
(Fig. 1) (Sakaguchi et al., 1984; White et al., 1984; Oeda et al., 1985;
Gonzalez et al., 1988a,b; Umeno et al., 1988; Larson et al., 1991b;
Khan et al., 2002; Wester et al., 2002; Yu et al., 2002; Scott et al.,
2003; Williams et al., 2003, 2004; Schoch et al., 2004) (Table 1). Al-
though some P450 amino acid sequences had been elucidated through
methods discussed earlier, the development of cloning, nucleotide se-
quencing, and DNA manipulation technologies would equip researchers
with powerful tools to generate libraries of P450 cDNA sequences
(Poulos et al., 1987; Gonzalez and Gelboin, 1992; Poulos and Follmer,
2022). Complementary DNA sequences for human P450 enzymes
were generated using mature mRNA from liver and lung tissue lys-
esates (Gonzalez and Gelboin, 1992). Purified mRNA was then used
to generate cDNA strands via reverse transcriptase, which then
underwent amplification by polymerase chain reaction to generate
double-stranded DNA fragments (Patil et al., 2022). These gene seg-
ments were then ready be digested with complementaryendonu-
clases and ligated into a vector for transformation of the amplified
cDNA into a suitable host organism (Unger et al., 1986; Gonzalez
and Gelboin, 1992; Patil et al., 2022). Due to observed sequence ho-

ology between species, human P450 sequences could be identified from
these libraries through the use of cDNA probes and antibodies
to rodent forms of P450 (Gonzalez et al., 1988b; Gonzalez and
Gelboin, 1992). Once cDNA sequences of interest were identified and
cloned using this approach, researchers set out to optimize the expres-
sion of these recombinant gene products in model organisms
so their catalytic activities, substrate specificities, and structure-
function relationships could be characterized (White et al., 1984;
Oeda et al., 1985; Gonzalez et al., 1988a,b; Gonzalez and Gelboin,
1992; Schoch et al., 2004; Guengerich, 2020).

One important example that used cDNA-directed expression to illu-
minate P450 structure was the investigation of rabbit microsomal P450
and its cotranslational insertion into rough microsomes from dog
pancreas (Sakaguchi et al., 1984, 1987) (Table 1). This would serve as an
effective model of mammalian P450 integration into membranes of
the endoplasmic reticulum, allowing researchers to interrogate the role of
key structural features responsible for initiating cotranslational insertion
and for conferring membrane solubility. Initial research demonstrated
that cotranslational insertion of P450 could not occur without the
presence of a ribonucleoprotein complex referred to as the signal
recognition particle (SRP) (Sakaguchi et al., 1984). However, fur-
ther investigation was required to identify the way in which micro-
sonal P450 recognizes the SRP to influence P450 function and
membrane topology. Hybrid cDNAs were generated to express chimer-
ic forms of the P450 protein for study, with each mutant posses-
sing unique modifications impacting the length and identity of the
N-terminal amino acid chain (Sakaguchi et al., 1987). Of the
several chimeric P450 proteins generated, only those possessing an
NH2 terminus at least 29 AAs in length were able to cotranslational-
ly insert into the membrane (Sakaguchi et al., 1987). This demonstrated
that the insertion signal that enables SRP-mediated cotranslation to the
membrane was contained within this 29 AA segment of the enzyme’s
NH2 terminus (Sakaguchi et al., 1987). This research broadened the un-
derstanding of the NH2 terminus present in eukaryotic P450 forms,
demonstrating the impact that NH2 truncation has on their topogenic
function and membrane solubility (Sakaguchi et al., 1987).

Interestingly, recombinant research efforts expressing rat P450mc in
*Saccharomyces cerevisiae* also generated a truncated form of the enzyme
that lacked an extended NH2 terminus (Oeda et al., 1985). Spectral and
catalytic analysis revealed almost identical properties between the trunc-
cated and nontruncated forms of the enzyme, demonstrating that this
NH2 terminus is not required for P450 catalytic activity (Oeda et al.,
1985) (Table 1). This knowledge would be widely applied to solubilize
mammalian P450 through deletion of the NH2 terminus, allowing re-
searchers to more easily purify and crystallize mammalian P450 enzymes
for structural and functional characterization (Larson et al., 1991b;
Hsu et al., 1993; Richardson et al., 1995; von Wachenfeldt et al.,
1997; Williams et al., 2000; Podust et al., 2001; Wester et al., 2002,
2003; Scott et al., 2003; Williams et al., 2003, 2004; Schoch et al.,
2004; Rowland et al., 2006; Hsu et al., 2018). This cDNA-directed approach
would soon be used to express truncated rabbit CYP2E1 in *E. coli* using
a pKKHC vector, an optimized derivative of the pBR322 plasmid dis-
cussed earlier (Larson et al., 1991a,b) (Table 1). These research efforts
further supported the notion that the NH2 terminus is not necessary
for retaining P450 monooxygenase function, with recombinant trunc-
cated CYP2E1 possessing identical spectral and catalytic properties
compared with full-length CYP2E1 isolated from rabbit liver mi-
crosomes (Larson et al., 1991a,b).

Similar success was also achieved in the recombinant expression of
rabbit CYP2C3 as a soluble dimer in *E. coli*. The CYP2C3 gene
insert with a modified NH2 terminus was digested with endonucleases
NdeI and EcoRi and transformed into an optimized pCWOri+ vector
that had been linearized with the same enzymes (von Wachenfeldt et al.,
1997) (Fig. 2). This general procedure is highlighted in Fig. 2, serving
as an example of heterologous P450 expression in *E. coli*. Additionally,
the key components of this vector that allow for strong transgene expres-
sion and inducible control are summarized in Fig. 2. This pCWOri+ de-

erived expression vector would also later prove useful in the recombinant
study of CYP2C5 (von Wachenfeldt et al., 1997; Wester et al., 2002,
2003).

Not long afterward, two mammalian P450 enzyme crystal structures
(CYP2B4, CYP2C5) were identified through recombinant expression in
*E. coli* (Wester et al., 2002, 2003; Scott et al., 2003) (Table 1). Both en-
zymes were expressed in soluble form through truncation of the NH2
terminus, allowing for their crystallization and structural characteriza-
tion. Crystallization in the presence and in the absence of substrate al-
lowed investigators to examine conformational differences that can
determine substrate recognition and catalytic function (Wester et al.,
2003). These experiments were also one of the first to add a histidine
tag to the C terminus of the recombinant protein to aid the purification
process (von Wachenfeldt et al., 1997). This strategy would also be em-
ployed in the study of human CYP2C8, which was uniquely expressed
in E. coli as stable homodimer complexed by two palmitic acid molecules (Schoch et al., 2004). These peripheral fatty acid binding sites can alter characteristics of the active site and have potential to modulate substrate binding as well as contribute to DDIs (Schoch et al., 2004) (Table 1). This observed dimerization of CYP2C8 in addition to its significantly larger active site cavity suggest a distinct metabolic function of this enzyme when compared with previously elucidated P450s. This makes sense when considering that the primary substrates of CYP2C8 are comparatively larger molecules such as taxol (Schoch et al., 2004). Further research crystallized CYP2C8 with a variety of different substrates to examine the role of protein flexibility and ligand-induced conformational changes in accomplishing substrate oxidation (Schoch et al., 2008).

The crystal structure of human CYP2C9 was also determined through heterologous expression of a NH2 truncated mutant, which revealed a similarly large active site capable of accommodating multiple warfarin molecules or an additional substrate of similar size (Williams et al., 2003) (Table 1). Given that CYP2C9 can be activated by some exogenous compounds including warfarin, it was hypothesized that the binding of some ligands may serve an allosteric function that promotes enzyme activity (Williams et al., 2003). Compared with CYP2C9, the crystal structure of recombinantly expressed human CYP3A4 revealed a slightly smaller active site as well as a characteristic cluster of phenylalanine residues above it, some of which are known to cooperate in enzyme selectivity and function (Khan et al., 2002; Williams et al., 2004). X-ray crystallography data were generated for CYP3A4 in the absence of ligand, in the presence of an inhibitor metyrapone, and with its substrate progesterone (Williams et al., 2004) (Table 1).

To determine how the properties and structures of CYP3A4 and CYP3A5 differ, both enzymes were recombinantly expressed in E. coli and complexed with a high-affinity substrate ronivar (Hsu et al., 2018). Crystallography data for these enzymes revealed differences in the shape and plasticity of their active sites upon substrate binding (Hsu et al., 2018) (Table 1). This structural data may reflect notable differences in the substrate dynamics and catalytic accessibility of these enzymes, potentially leading to divergent metabolic pathways or varying enzymatic efficiencies for shared metabolic pathways (Hsu et al., 2018). Additionally, enzymes of the CYP3A family can be differentially expressed across ethnic populations, with allelic variations having a significant effect on enzyme abundance and metabolic capability (Hirota et al., 2004; Yamaori et al., 2005). Given that the CYP3A family is the largest single contributor to drug metabolism, illuminating the structural features and processes that govern their catalytic activity and specificity is critically important in designing safe therapies, predicting DDIs, and understanding individual differences in drug metabolism (Zanger et al., 2008).

One recent advancement in the structural understanding of this enzyme family is the solved crystal structure of neonatal CYP3A7, a critical enzyme in hormone homeostasis and pediatric drug metabolism (Li and Lampe, 2019; Sevrioukova, 2021). Structural characterization of CYP3A7 was previously lacking, as the enzyme proved difficult to crystallize in its wild-type form. Using site-directed mutagenesis, a recombinant CYP3A7 variant capable of crystallization was generated to allow for structural characterization and comparison with other CYP3A enzymes (Sevrioukova, 2021) (Table 1). In many ways rDNA technologies have provided unprecedented insight into the catalytic structure, conformational dynamics, and substrate specificities of the most integral drug-metabolizing enzymes. Even today we continue to build upon our understanding of P450 monooxygenases, with recent investigations providing additional context regarding the nature of active site plasticity and redox partner selectivity (Poulos and Follmer, 2022).

Another cytochrome P450 enzyme that performs a significant share of xenobiotic biotransformations is human CYP2D6, as it is estimated to metabolize up to 20% to 30% of commonly prescribed drugs. The crystal structure was solved once again through heterogenous expression of a soluble recombinant CYP2D6 in E. coli utilizing a pCWori + derived vector (Rowland et al., 2006) (Table 1). This experiment revealed a structure most closely resembling that of CYP2C9, and the crystallography data enabled further research efforts toward building a more detailed understanding of the structural basis of ligand binding and specificity (Rowland et al., 2006). This information proved especially critical in understanding the many biotransformations mediated by CYP2D6 and CYP2C9, as their allelic variants between ethnic groups and polymorphisms between individuals can have a dramatic effect on metabolic activity, drug disposition, and therapeutic outcomes (Higashi et al., 2002; Yu et al., 2002; Ingelman-Sundberg, 2005; Zhou, 2009a,b).

The rDNA platform has been adapted to interrogate the impact of specific CYP polymorphisms present in the population, uncovering mechanistic rationale for unique phenotypes of drug metabolism. The ability to recombinantly express specific P450 polymorphisms has allowed researchers to quantify enzymatic differences in the presence of any substrate, providing critical knowledge in our understanding of pharmacogenomic variability between individuals and populations. This knowledge has proven invaluable in drug design and clinical practice, as CYP polymorphisms and genetic variants can result in altered ADME processes, DDIs, and unexpected adverse effects. One key example of this is observed in warfarin-treated individuals with PYP2C9 polymorphisms, who are at increased risk of major bleeding events and excessive anticoagulation associated with increased drug exposure (Higashi et al., 2002). CYP2C9 polymorphisms have also been shown to promote genotype-dependent DDIs when warfarin or other major substrates are coadministered with an inducer or inhibitor of CYP2C9 (Higashi et al., 2002; Gardin et al., 2019; Cheng et al., 2022). Research utilizing rDNA technologies has revealed many genotype-specific differences in PK function and individual drug disposition, providing understanding that is critical to rational drug design, clinical dose management, and DDI prediction. Elucidating these pharmacogenomic differences has been key in developing and maintaining safe therapies, and this knowledge will prove increasingly useful as pharmacogenomic testing of patients becomes more commonplace and cost effective (Goh et al., 2017; Sayer et al., 2021).

In addition to providing valuable insight into the most predominant drug metabolizing P450s, rDNA technologies have also been used to characterize P450 enzymes that biosynthesize critical endogenous compounds. One predominant example of this is human CYP17A1, which is known to catalyze the synthesis of androgens and other important steroids (Miller and Aichus, 2011). Because of this, CYP17A1 inhibition represents an attractive strategy to prevent androgen synthesis and treat castration-resistant prostate cancer. The crystal structures of CYP17A1 bound to two inhibitors were determined by recombinant expression of a truncated version in E. coli (DeVore and Scott, 2012) (Table 1). These initial findings elucidated the structural features integral to CYP17A1’s catalytic activity, further enhanced soon afterward by crystal structures of CYP17A1 with all four of its major endogenous substrates and clinically relevant inhibitors (Petrunak et al., 2014, 2017). Structural characterization of CYP17A1 enabled by recombinant expression provided critical information in the rational drug design of CYP17A1 inhibitors. Additionally, rDNA technologies allowed investigators to develop more selective CYP17A1 inhibitors through assessment of off-target interactions with other steroidogenic P450 enzymes such as CYP21A2 (Fehl et al., 2018).
Another enzyme that mediates critical biosynthetic reactions is cytochrome P450 aromatase (CYP19A1), which is known to synthesize estrogens from androgens. While the crystal structure of CYP19A1 was first determined using aromatase purified from human placenta, subsequent research efforts recombinantly expressed truncated CYP19A1 to further its structural characterization and interrogate the functional roles of specific amino acid residues (Ghosh et al., 2009, 2010; Lo et al., 2013) (Table 1). This insight into the structure of CYP19A1 and its androgen-specific active site aided the intelligent design of aromatase inhibitor drugs used to treat hormone receptor-positive breast cancers (Ghosh et al., 2012). Indeed, this same approach has proven similarly useful in the characterization of many P450 enzymes that catalyze the biosynthesis of important endogenous compounds (Brixius-Anderko and Scott, 2019, 2021; Liu et al., 2022). These investigations have supported the rational design of selective inhibitor drugs, with recent examples being the characterizations of P450 11B1, P450 11B2, and P450 8B1, which catalyze the production of cortisol, aldosterone, and bile acid, respectively (Brixius-Anderko and Scott, 2019, 2021; Liu et al., 2022) (Table 1).

Recombinant RNAs for Studying ADME Gene Regulation

New advancements in epigenetics in DMPK as well as ADME gene regulation and function are critical in designing the next generation of safe and effective therapeutics. One of the most important factors in the posttranscriptional regulation of ADME genes are the aforementioned miRNAs, whose wide range of regulatory targets and functions allow them to drive therapy resistance or sensitivity depending on the nature of miRNA, disease, and context (Liu et al., 2015; Mognoto and Celotti, 2015; O’Brien et al., 2018; Li et al., 2019). As discussed earlier, successes in recombinant technologies have also provided scientists with the tools to produce biologic RNA molecules as research tools and potential therapeutics. To this end, the Yu laboratory has established a novel approach to generate recombinant or bioengineered RNA (BioRNA) molecules in high yield utilizing E. coli fermentation (Li et al., 2021; Tu et al., 2021) (Fig. 3). These BioRNAs synthesized and processed inside living cells more accurately represent the properties and chemical modifications of endogenous miRNAs compared with commercially available chemical mimics (Ho and Yu, 2016; Yu et al., 2019). This approach has been used to generate a collection of humanized miRNAs more representative of their natural state, allowing for more accurate interrogation of their epigenetic regulation of ADME genes to influence drug disposition and therapeutic outcomes (Yu et al., 2019; Tu et al., 2021).

To bioengineer these RNA molecules in E. coli, the pBSTNAV vector is used to accomplish cloning and expression. This vector contains an ampicillin resistance gene, compatible ColE1 and pBR322 ori sites, a strong lipoprotein promoter, and a ribosomal RNA operon termination sequence (rmfC) (Pouchon et al., 2009; Tu et al., 2021). These features allow for antibiotic-based selection of positive transformants and enable cloning and expression of a noncoding RNA (ncRNA) segment within the plasmid (Scott, 1985). This is accomplished using a gene insert encoding a fused transfer RNA (tRNA) pre-miRNA carrier molecule that can be stably expressed in E. coli (Li et al., 2021) (Fig. 3). The identity of the tRNA and pre-miRNA sequences influence expression efficiency, and research efforts investigating different tRNA scaffolds and pre-miRNA species have resulted in an optimized stable carrier system capable of consistently expressing BioRNAs at levels surpassing 30% of total bacterial RNA (Ho et al., 2018; Tu et al., 2021). This stable hybrid carrier has been refined to achieve production at high yield, with human tRNA fused hsa-pre-miR-34a demonstrating consistently high expression with many different ncRNA cargos (Li et al., 2021). Within a stable human tRNA/hsa-pre-miR34a carrier, the miR-34a duplex can be replaced with payload miRNA sequences of interest to generate recombinant miRNA agents (Fig. 3) (Tu et al., 2021). This duplex can also be replaced with other small RNAs such as siRNA or even multiple payload RNA sequences, highlighting the flexibility of this approach (Petrek et al., 2019).

Through these research efforts, the Yu laboratory has developed a robust bioengineering platform capable of producing humanized BioRNAs in high yield to investigate their role in the posttranscriptional regulation of ADME genes as well as their impact on therapeutic susceptibility and treatment outcomes. Websites such as miRBase, miRDB, and Targetscan are valuable resources in identifying ncRNA segments of interest as well as investigating gene transcripts that are predicted regulatory targets of specific miRNA sequences. Sequences encoding target BioRNAs are modified with appropriate cloning sites for vector insertion and amplified by polymerase chain reaction utilizing forward and reverse cloning primers with 15 nucleotide overlaps (Li et al., 2021; Tu et al., 2021). Amplified BioRNA segments are then ligated into the pBSTNAV plasmid following linearization with the restriction enzymes EcoRI and PstI, with strategies utilizing EagI and SacII being used to some success in the past (Li et al., 2021; Tu et al., 2021) (Fig. 3). Plasmids containing target BioRNA inserts are confirmed by sequencing analyses and used to transform HST08 competent E. coli cells to achieve overexpression of rRNAs (Li et al., 2021; Tu et al., 2021). Total RNA is then isolated from bacterial culture, and successful expression of target BioRNA is verified through the appearance of a strong band of expected size on an acrylamide-urea gel (Tu et al., 2021). Anion exchange fast protein liquid chromatography is used to separate the RNAs, allowing the collection of fractions containing the target BioRNA (Tu et al., 2021). These fractions can be run on a gel to determine their purity in a semiquantitative manner, with the purest of these being pooled together to undergo desalting and concentration procedures. These isolated BioRNAs then have their purity quantitatively assessed by high performance liquid chromatography, and the level of bacterial endotoxin within each is determined (Tu et al., 2021).

The streamlined procedure outlined here in brief has been used to express and purify a collection of humanized BioRNA agents for use in research. These BioRNAs serve as prodrugs for the warhead miRNAs they contain, being processed by endogenous machinery to generate high levels of target miRNAs as detectable by stem-loop reverse transcription-quantitative polymerase chain reaction or RNA sequencing (Li et al., 2021). Selective processing of bioengineered miRNA prodrugs to their target miRNAs has been demonstrated in several in vitro and mouse models, as well as their ability to regulate gene expression to impact cellular processes and chemosensitivities (Ho et al., 2018; Jilek et al., 2019; Li et al., 2021) (Table 2). This innovative platform has been leveraged to investigate the therapeutic potential and epigenetic role of many small RNAs across several disease states and target protein types. In the remainder of this review, we highlight several recent studies utilizing these strategies to interrogate miRNA’s role in the posttranscriptional regulation of gene products integral to drug metabolism and disposition (Table 2).

The first investigation utilizing recombinant miRNAs to modulate ADME gene expression was published in 2014, and this study only used a tRNA scaffold to achieve expression of hsa-pre-miR-27b (Li et al., 2014). Previous research had shown miR-27b to be a negative regulator of CYP3A4 and vitamin D receptor (VDR) through targeting the 3’-untranslated regions of these transcripts (Pan et al., 2009), making miR-27b an ideal choice to test the effectiveness of recombinant miRNA molecules in accomplishing target gene regulation. The VDR is a nuclear receptor that is known to play a critical role in the transcriptional control of CYP3A4 (Qin and Wang, 2019). The tRNA fused to hsa-pre-miR-27b was successfully expressed in E. coli but demonstrated low overall
abundance with recombinant RNA making up only 2% of total bacterial RNA (Li et al., 2014). Despite this challenge, tRNA/hsa-pre-miR-27b was successfully purified from bacteria and transfected into human cells for further study. The recombinant RNA was successfully processed to mature miR-27b in LS180 cells, and further analysis demonstrated the suppression of CYP3A4 and VDR protein expression as well as mRNA transcript abundance (Li et al., 2014). This reduced protein expression also corresponded with a decrease in CYP3A4 enzymatic activity.

Fig. 3. Design of recombinant RNA expression plasmids for fermentation production in E. coli. (A) Recombinant or BioRNA agents are expressed with hybrid tRNA/pre-miRNA carriers, protecting them from RNase degradation and generating consistently high yield of desired BioRNA from bacterial culture. The miR-34a duplex highlighted in red can be replaced with a payload ncRNA of interest to be recombinantly expressed as part of this hybrid carrier. (B) An expression vector derived from the pBSTNAV plasmid is used for recombinant BioRNA expression. This vector contains a strong lipoprotein gene promoter as well as a ribosomal RNA operon transcription terminator (rrnC) to ensure efficient expression of the BioRNA insert. This vector also contains ColE1 and F1 ori sites, which are recognized by E. coli and M13 bacteriophages, respectively. The gene encoding ampicillin resistance (AmpR) is also present in the vector, which enables positive selection of positive transformants through antibiotic treatment. (C) Payload ncRNA is identified and the resulting BioRNA sequence is defined accordingly. Restriction sites are added to the ends of the BioRNA coding sequence, and cloning primers are designed to replace mature miR-34a-5p and its complementary strand with the desired payload ncRNA. This BioRNA insert is then amplified by polymerase chain reaction in preparation for vector insertion. The vector is then linearized by endonucleases EcoRI and PstI, allowing insertion and ligation of the recombinant BioRNA gene insert. Once verified, the BioRNA expression plasmid is then ready for transformation into HST08 E. coli for recombinant BioRNA expression.
TABLE 2
Modulation of ADME gene expression as well as drug metabolism and transport by recombinant RNA agents

<table>
<thead>
<tr>
<th>BioRNA</th>
<th>ADME gene</th>
<th>Effects</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-27b-3p</td>
<td>CYP3A4</td>
<td>Downregulates CYP3A4 mRNA transcript abundance, protein expression, and 1'-hydroxylase activity in LS180 cells</td>
<td>Li et al., 2014; Li et al., 2019</td>
</tr>
<tr>
<td>miR-1291-5p</td>
<td>ABCC1</td>
<td>Downregulates ABCC1 protein expression in MCF-7 and PANC-1 cells and sensitizes them to doxorubicin treatment</td>
<td>Li et al., 2015</td>
</tr>
<tr>
<td>miR-124-3p</td>
<td>ABCC4</td>
<td>Downregulates ABCC4 and STAT3 protein expression in A549</td>
<td>Ho et al., 2018</td>
</tr>
<tr>
<td>miR-328-3p</td>
<td>ABCC5/ABCC4</td>
<td>Downregulates ABCG2/BCRP mRNA transcript abundance and protein expression in MCF-7/MX100 cells to increase intracellular accumulation of mitoxantrone and sensitize cells to treatment</td>
<td>Li et al., 2019</td>
</tr>
<tr>
<td>let-7c-5p</td>
<td>ABCC5/ABCC4</td>
<td>Downregulates ABCG5 and ABCC4 protein expression in Huh7 cells, promotes intracellular accumulation of 5-FU, and exhibits synergistic effects inhibiting Huh7 cell viability</td>
<td>Jilek et al., 2020</td>
</tr>
</tbody>
</table>

BCRP, breast cancer resistance protein; 5-FU, fluorouracil.

Conclusions and Perspectives

tRNA technologies have revolutionized the field of drug metabolism and PK, equipping investigators with the tools to produce pure ADME gene products for research and development. Advancements in our understanding and utilization of expression vectors and host organisms has provided researchers with a versatile and adaptable toolbox to express and purify almost any gene of interest. This strategy has been widely employed in the study of P450 monoxygenases, where heterologous expression of these drug-metabolizing enzymes has allowed investigators to examine their conformational interactions with xenobiotics and endogenous compounds. In this way, recombinant technologies have provided novel insights into CYP enzymes and other ADME gene products, catapulting forward the fields of DMPK. These research efforts have armed investigators with critical understanding of the molecular mechanisms underlying xenobiotic biotransformation and disposition, providing information that is invaluable to the development of safe and effective therapeutics. Recombinant protein expression continues to hold incredible potential for future DMPK research, improving availability of robust research tools in addition to aiding the design of new antibodies and specific inhibitor drugs by providing facile means to express and purify specific transgenes of interest.

As recombinant technologies have granted us further insight into the gene products behind DMPK processes, they have also revealed an

demonstrated by a reduced capacity to perform the 1'-hydroxylation of probe drug midazolam (Li et al., 2014). In 2015, this same strategy was used to produce recombinant tRNA/hsa-pre-miR-1291 for functional study in human carcinoma cells (Li et al., 2015). This chimeric RNA was processed to mature miR-1291 in MCF-7 and PANC-1 cell lines, where it was able to effectively downregulate the protein expression of target ABCC1 (Li et al., 2015) (Table 2). Overexpression of ABCC1 and similar transporters is a common occurrence in cancers, which is believed to contribute to therapeutic evasion and multidrug resistance. Introduction of recombinant tRNA/hsa-pre-miR-1291 into PANC-1 cells also enhanced their sensitivity to doxorubicin, a known substrate of the ABCC1 transporter (Li et al., 2015). These results highlighted the early potential of recombinant miRNAs in modulating ADME gene expression to alter drug exposure, overcome resistance, and improve therapeutic outcomes.

The identification and use of a stable tRNA/pre-miRNA carrier to produce functional small RNA agents was first reported in 2015, as most target RNAs cannot be heterogeneously expressed with a tRNA scaffold alone (Chen et al., 2015). A few years later in 2018, another research paper was published that optimized the use of the discussed hybrid tRNA/pre-miRNA carrier to achieve substantially higher yields of target ncRNA sequences from bacterial fermentation, with a greater success rate (Ho et al., 2018). Utilizing this approach, recombinant miR-124 was generated and used to transfect human cells. The administered agent was then selectively processed to miR-124-3p within the cells, effectively downregulating its established targets ABCC4, as well as other cancer-related genes, and inhibiting proliferation of carcinoma cells (Ho et al., 2018) (Table 2). Further research in 2019 used these innovations to produce recombinant tRNA/hsa-pre-miR-1291 for functional study in human carcinoma cells (Li et al., 2019) (Table 2). Once administered to LS180 cells, BioRNA/miR-27b-3p was processed to mature miR-27b-3p and again successfully suppressed expression of its targets CYP3A4 and VDR. This once again corresponded with a reduction in enzyme activity as demonstrated by a decrease in midazolam 1’-hydroxylase metabolism (Li et al., 2019). This study also produced BioRNA/miR-328-3p for functional study, and it was selectively processed to mature miR-328-3p in human lung carcinoma cells. MiR-328-3p is a known regulatory miRNA for ABCG2, which is a critical drug transporter implicated in chemotherapeutic resistance that is also commonly known as breast cancer resistance protein (Li et al., 2011). Transfection with BioRNA/miR-328-3p resulted in knockdown of ABCG2 protein levels and mRNA transcripts, sensitizing chemo-resistant MCF7/MX100 cells to mitoxantrone treatment via promoting the drug’s intracellular accumulation (Li et al., 2019) (Table 2). Mitoxantrone is an established substrate of ABCG2, providing an explanation as to why downregulation of this transporter could sensitize cancer cells to treatment. Intracellular accumulation of the drug was assessed by examining mitoxantrone fluorescence intensity, which was higher across BioRNA/miR-328-3p treated cells (Li et al., 2019).

Another study published in 2020 investigated the role of epigenetic role of BioRNA/let-7c-5p, an agent with demonstrated promise in inhibiting hepatocellular carcinoma growth (Jilek et al., 2019, 2020). This BioRNA was processed to mature let-7c-5p within living cells, effectively suppressing the protein levels of ABCC4 and ABCC5 transporters (Jilek et al., 2020) (Table 2). Downregulation of these transporters by BioRNA/let-7c-5p promoted intracellular accumulation of cancer drug 5-FU, explaining the synergistic effects observed with these agents in the inhibition of cancer cell viability (Jilek et al., 2020). While most ADME research utilizing this platform so far has been centered on miRNAs, it has also been used to elucidate the epigenetic role of nuclear factor erythroid 2-related factor 2 (NRF2) through a corresponding siRNA. NRF2 is a transcription factor understood to play a critical role in the induction of ADME gene products such as ABC transporters, with NRF2 overexpression being a commonly observed phenomenon in multidrug resistant cancers (Lu et al., 2016; Li et al., 2018; Hu et al., 2019). Administration of BioRNA/NRF2-siRNA resulted in significant downregulation of the transporters ABCC3, ABCC4, and ABCC5, sensitizing human cancer cells to commonly employed chemotherapeutics doxorubicin, cisplatin, and sorafenib (Li et al., 2018).

Conclusions and Perspectives

rDNA technologies have revolutionized the field of drug metabolism and PK, equipping investigators with the tools to produce pure ADME gene products for research and development. Advancements in our understanding and utilization of expression vectors and host organisms has provided researchers with a versatile and adaptable toolbox to express and purify almost any gene of interest. This strategy has been widely employed in the study of P450 monoxygenases, where heterologous expression of these drug-metabolizing enzymes has allowed investigators to examine their conformational interactions with xenobiotics and endogenous compounds. In this way, recombinant technologies have provided novel insights into CYP enzymes and other ADME gene products, catapulting forward the fields of DMPK. These research efforts have armed investigators with critical understanding of the molecular mechanisms underlying xenobiotic biotransformation and disposition, providing information that is invaluable to the development of safe and effective therapeutics. Recombinant protein expression continues to hold incredible potential for future DMPK research, improving availability of robust research tools in addition to aiding the design of new antibodies and specific inhibitor drugs by providing facile means to express and purify specific transgenes of interest.

As recombinant technologies have granted us further insight into the gene products behind DMPK processes, they have also revealed an
increasingly complex picture of drug metabolism that can vary widely between populations and individuals. Keeping this in mind, it is essential that investigators employ RNA technologies in new and innovative ways to interrogate the function and regulation of ADME gene products. Recently, miRNAs have emerged as important factors in this equation, with their roles in posttranscriptional regulation of ADME genes having a significant impact on drug metabolism and transport as well as PK and efficacy (Yu et al., 2016; Jilek et al., 2017).

To build upon our understanding of miRNA-directed regulation and address the widespread usage of synthetic miRNA mimics in research, an innovative bioengineering platform has been developed, capable of generating tens milligram quantities of pure recombinant RNA agents from only a liter of bacterial culture. These humanized BioRNA agents are produced and processed inside living cells, serving as a more accurate representation of natural endogenous miRNAs. These BioRNAs do not require added chemical modifications to maintain their intracellular stability. It has been proven these biologic RNA agents are selectively processed to the target miRNAs or siRNAs in human cells to selectively regulate target gene expression. Additionally, the hybrid RNApre-miRNA carrier used in this technology can achieve high levels of expression with a wide variety of small RNAs (Yu and Tu, 2022; Traber and Yu, 2023). Using this novel platform, the Yu laboratory has amassed a collection of humanized BioRNA agents so that their utilities in modulating ADME gene expression may be interrogated. This novel bioengineering strategy has been successfully used in the study of several ADME genes, effectively regulating expression of targets such as P450 enzymes and ABC transporters to alter drug metabolism, transport, exposure, and therapeutic outcome. The BioRNA platform has greatly increased the quality, availability, and adaptability of miRNA research tools, making it a unique and valuable addition to DMPK research as well as RNA technologies as a whole.

Authorship Contributions

Performed data analyses: Cronin, Yu.
Wrote and contributed to the writing and revision of the manuscript: Cronin, Yu.

References


Downloaded from dl-id:aspetjournals.org at ASPET Journals on December 29, 2023
Lederberg J (1952) Cell genetics and hereditary symbiosis.


Maniatis T, Fritsch EF, and Sambrook J (1983) Molecular Cloning, Cold Spring Harbor Laboratory, 


