P-glycoprotein structure and evolutionary homologies

Irene Bosch¹ & James M. Croop²

¹ The Dana-Faber Cancer Institute, Harvard Medical School, Boston, MA, U.S.A.

² The Section of Pediatric Hematology/Oncology, James Whitcomb Riley Hospital for Children, Indiana University, Indianapolis, IN 46202, U.S.A.

Received 25 May 1998; accepted 25 May 1998

Key words: ABC transporters, MDR, multidrug resistance, P-glycoprotein

Abstract

Analysis of multidrug resistant cell lines has led to the identification of the P-glycoprotein multigene family. Two of the three classes of mammalian P-glycoproteins have the ability to confer cellular resistance to a broad range of structurally and functionally diverse cytotoxic agents. P-glycoproteins are integral membrane glycoproteins comprised of two similar halves, each consisting of six membrane spanning domains followed by a cytoplasmic domain which includes a nucleotide binding fold. The P-glycoprotein is a member of a large superfamily of transport proteins which utilize ATP to translocate a wide range of substrates across biological membranes. This superfamily includes transport complexes comprised of multicomponent systems, half P-glycoproteins and P-glycoprotein-like homologs which appear to require approximately 12 α -helical transmembrane domains and two nucleotide binding folds for substrate transport. P-glycoprotein homologs have been isolated and characterized from a wide range of species. Amino acid sequences, the similarities between the halves and intron/exon boundaries have been compared to understand the evolutionary origins of the P-glycoprotein.

Introduction

Resistance to chemotherapeutic agents remains the major obstacle in the successful therapy of human cancer. Many tumors are intrinsically resistant to many of the most potent cytotoxic agents used in cancer therapy. Other tumors, initially sensitive, recur and are resistant not only to the initial therapeutic agents but also to other drugs which the tumor has not previously been exposed. This resistance of human tumors to multiple chemotherapeutic agents has been difficult to study in the clinical setting. However, significant progress has been made in characterizing the phenomenon of resistance to multiple cytotoxic agents *in vitro* (Bosch and Croop, 1996; Endicott and Ling, 1989; Gottesman and Pastan, 1989; Roninson, 1991). Cell lines have been selected for resistance to anticancer agents, many of which are antibiotics originating in microorganisms or alkaloids from plants, by slowly increasing the concentration of the cytotoxic agent in a step-wise fashion. Derivative lines, many thousand fold resistant to the selecting agent, have been

generated which are cross resistant to a broad spectrum of structurally and functionally diverse cytotoxic compounds. These multidrug resistant cell lines have provided an experimental model for the analysis of cellular resistance to multiple cytotoxic agents. An understanding of the biological basis of multidrug resistance has emerged which points to several distinct genetic mechanisms responsible for cross resistance to specific cytotoxic agents. One of the first and most potent mechanisms identified, overexpression of P-glycoprotein, has become an attractive target to improve anti-cancer therapy. P-glycoprotein expression has also provided a selectable marker for manipulating hematopoietic cells in gene therapy protocols.

Multidrug resistant cell lines were initially generated using anthracyclines, colchicine, vinca alkaloids, or actinomycin D (Nielsen and Skovsgaard, 1992). The cell lines usually displayed the highest level of resistance to the selecting agent with lesser degrees of resistance to the other cytotoxic agents in these categories (Table I). Although the compounds appear quite

Table 1. Compounds included in the multidrug resistant phenotype or which the P-glycoprotein transports

Agents included in the multidrug resistance phenotype	
Doxorubicin	Actinomycin D
Daunorubicin	Puromycin
Mitoxantrone	Emetine
Colchicine	Ethidium Bromide
Vinblastine	Gramacidin D
Vincristine	Mithramycin
Taxol	Lipophilic Cations
Etoposide	Lipophilic Peptides
Rhodamine 123	^{99m} Tc-Sestamibi
Verapamil	Phospholipids

diverse in their structure and mode of action, many are natural products and can usually be categorized as small, lipophilic cations (Beck, 1990). In addition, a broad range of compounds, which often share structural characteristics with these substrates, have been shown to be capable of reversing multidrug resistance (Ford and Hait, 1990).

Several initial observations provided clues for understanding the mechanism of P-glycoprotein mediated multidrug resistance. First, multidrug resistant cell lines accumulated less of the cytotoxic agents than non-resistant lines. The decreased accumulation was associated with an energy dependent outward efflux of the compounds (Dano, 1973; Skovsgaard, 1978). Second, Juliano and Ling (Juliano and Ling, 1976) identified a 170 kD membrane glycoprotein, termed P-glycoprotein (Pgp), overexpressed in multidrug resistant cell lines. The amount of P-glycoprotein in the plasma membrane appeared to correlate with the level of drug resistance. Finally, karyotypic analysis of multidrug resistant cell lines indicated the presence of double minute chromosomes or homogeneously staining regions, abnormalities associated with gene amplification (Biedler and Riehm, 1970). These observations, consistent with overexpression of a gene product responsible for decreased accumulation of cytotoxic agents, directed the means for elucidating the mechanism of multidrug resistance.

Isolation of cDNAs encoding the mammalian P-glycoprotein

Using a variety of strategies, genes overexpressed in multidrug resistant cell lines have been isolated

Table 2. Classification and nomenclature of the mammalian P-glycoproteins. Classes I and II convey multidrug resistance

Class	Mammalian P-glycoproteins		
	I	II	III
Human	<i>mdr1</i>		<i>mdr2/3</i>
Mouse	<i>mdr3</i> <i>mdr1a</i>	<i>mdr1</i> <i>mdr1b</i>	<i>mdr2</i> <i>mdr2</i>
Hamster	<i>pgp1</i>	<i>pgp2</i>	<i>pgp3</i>

and characterized. A series of cDNAs displaying a high level of nucleic acid similarity, each encoding a polypeptide with features of the P-glycoprotein have been identified. These cDNAs have been classified based on sequence similarity and have been designated members of the *mdr* or *pgp* multigene family (Table II).

Specific DNA sequences amplified in multidrug resistant cell lines were first identified using the technique of DNA renaturation in agarose gels (Roninson *et al.*, 1984). DNA sequences amplified in two independently derived hamster multidrug resistant cell lines were identified with a total length of at least 150 kbp. A 1.1 kbp BamHI fragment amplified in both cell lines was excised from the gel and cloned. This probe indicated that the 1.1 kbp BamHI fragment was amplified in both cell lines and that the level of drug resistance correlated with the level of amplification. The fragment did not, however, hybridize to a mRNA species. Using the 1.1 kbp BamHI fragment, a contiguous 140 kbp domain amplified in the two hamster multidrug resistant cell lines was isolated (Gros *et al.*, 1986c). Probes spanning the amplified domain indicated that a mRNA species of approximately 5 kbp was overexpressed in the multidrug resistant cell lines. The hamster probes were used to isolate two distinct cDNAs from a non-drug resistant, mouse pre-B cell library (Gros *et al.*, 1986a). One of the murine cDNAs, *mdr1*, was capable of conferring the complete multidrug resistant phenotype when transfected into drug sensitive cell lines (Croop *et al.*, 1987; Gros *et al.*, 1986a). The second cDNA, *mdr2*, although highly homologous to the first, did not convey drug resistance (Gros *et al.*, 1988). A third murine cDNA, *mdr3*, also capable of conveying multidrug resistance, was sub-

sequently isolated using probes from the previously cloned cDNAs (Devault and Gros, 1990).

Two distinct human genomic clones were isolated by cross hybridization with probes from the amplified hamster domain (Roninson *et al.*, 1986). One of these clones, human MDR1, identified amplified DNA sequences and a 4.5 kbp mRNA species overexpressed in multidrug resistant cell lines. MDR1 was subsequently used to screen a cDNA library derived from a multidrug resistant cell line to isolate a series of overlapping human cDNA clones which encoded sequences capable of conveying multidrug resistance (Chen *et al.*, 1986; Ueda *et al.*, 1987). Similarly, using an independently derived hamster *mdr* probe (Van der Bliek *et al.*, 1986), a novel set of cDNAs with high sequence similarity to human *mdr1* were isolated from a human liver cDNA library and designated *mdr3* (Van der Bliek *et al.*, 1987; van der Bliek *et al.*, 1988). As with mouse *mdr2*, transfection studies indicated that human *mdr3*, was not capable of conveying multidrug resistance.

Using the C219 monoclonal antibody directed against the P-glycoprotein (Kartner *et al.*, 1985), a 600 basepair clone, λ CHP1, was isolated from an expression library generated from a multidrug resistant hamster cell line (Riordan *et al.*, 1985). Molecular analysis of multidrug resistant cell lines with λ CHP1 indicated that the gene encoding the P-glycoprotein was amplified and a mRNA of approximately 4.7 kbp was overexpressed. Subsequently, three members of the hamster *pgp* multigene family have been isolated using clones derived from λ CHP1 (Endicott *et al.*, 1991).

Analysis of RNA transcripts overexpressed in multidrug resistant cell lines similarly identified cDNAs encoding P-glycoprotein. Differential screening of cDNA libraries from drug sensitive and drug resistant cell lines identified a series of 5 genes overexpressed in the multidrug resistant cell lines (Van der Bliek *et al.*, 1986). All five genes were present on a single amplicon. Only the cDNA which cross hybridized with λ CHP1, however, was consistently overexpressed in a variety of multidrug resistant cell lines (de Bruijn *et al.*, 1986). Similarly, using Cot fractionated genomic DNA to generate a probe from amplified sequences, cDNAs encoding the P-glycoprotein were identified in a library generated from a multidrug resistant cell line (Scotto *et al.*, 1986).

Cross hybridization (Gros *et al.*, 1986c; Ueda *et al.*, 1986) and sequence comparisons indicate that the cDNAs isolated from the human, murine and hamster

Figure 1. Predicted structure and membrane orientation of the P-glycoprotein. Cluster of glycosylation sites in first external domain and nucleotide binding folds identified.

libraries are members of a multigene family which encodes the P-glycoprotein. Comparisons of intron sequences (Ng *et al.*, 1989), untranslated sequences (Endicott *et al.*, 1987; Hsu *et al.*, 1989) and coding regions (Endicott *et al.*, 1991; Hsu *et al.*, 1989) have identified three classes of mammalian P-glycoproteins as indicated in Table II. The mammalian P-glycoprotein multigene families are clustered in tandem on a single chromosome, chromosome 7 in humans (Chin *et al.*, 1989; Lincke *et al.*, 1991; Raymond *et al.*, 1990). Classes I and II are capable of conveying multidrug resistance, while class III cannot. Systematic searches have never identified a human class II, suggesting that classes I and II represent a gene duplication occurring after the separation of murine and human speciation (Chin *et al.*, 1989; Ng *et al.*, 1989).

The predicted structure of the P-glycoprotein

The isolation of cDNAs encoding the P-glycoprotein enabled analysis of the structure of the polypeptide. The nucleic acid sequences of the cDNAs which encode the mammalian P-glycoproteins predict an integral membrane glycoprotein with remarkably similar structures (Figure 1). The length of the mammalian P-glycoproteins vary from 1276 to 1281 amino acids predicting a molecular weight of approximately 140 kD. This is consistent with purified or labelled P-glycoprotein which migrates at a molecular weight of 130–180 kD in polyacrylamide gels. Although P-glycoprotein mobility in polyacrylamide gels is sensitive to the conditions employed (Greenberger *et al.*, 1988b), approximately 10–15 kD of the observed molecular weight is accounted for by N-linked gly-

cosylation (Greenberger *et al.*, 1988a). Indeed, the predicted amino acid structure of the P-glycoprotein indicates 7–10 potential N-linked glycosylation sites [N-X-S/T].

Examination of the amino acid sequence identifies a number of significant structural features (Figures 1 and 2). Independent hydrophobicity analyses using the methods of Kyte-Doolittle (Gros *et al.*, 1986b) and Eisenberg (Chen *et al.*, 1986) predict a polypeptide which traverses the plasma membrane 12 times. A short, highly charged, cytoplasmic domain precedes three membrane loops which are followed by a large cytoplasmic domain. A cluster of potential N-linked glycosylation sites are consistently located in the first external loop in virtually all P-glycoproteins. The large cytoplasmic domain includes the paired consensus sequences G-x(4)-G-K-[ST] and [RK]-x(3)-G-x(3)-L-[Hydrophobic](3) separated by approximately 120 amino acids which function as a nucleotide binding fold in a wide range of ATPases (Walker *et al.*, 1982). The cytoplasmic domain is followed by three additional membrane loops and another large cytoplasmic domain with the consensus sequence for a second nucleotide binding fold. Experimental evidence indicates that the P-glycoprotein binds and hydrolyzes ATP (Ambudkar *et al.*, 1992; Sarkadi *et al.*, 1992) which is required for drug transport (Horio *et al.*, 1988).

This model of the P-glycoprotein suggests that it is comprised of two halves, each with a similar structure. A comparison of the amino acid sequence of the carboxy and amino halves indicates approximately 40% identical and an additional 25% conserved amino acid homology which has suggested duplication of an ancestral gene (Figure 3). The highest similarity between the moieties surrounds the nucleotide binding folds in the large cytoplasmic domains with almost 60% amino acid identity with an additional 20% conserved substitutions. When the two halves of the P-glycoprotein are aligned, a short segment termed the 'linker' region (Van der Bliek *et al.*, 1987) which does not readily align with the corresponding region in the proximal half of the polypeptide appears to bridge the amino and carboxy halves. This highly charged region is capable of self annealing *in vitro* and has been posed to play a role in the potential dimerization of P-glycoprotein (Juvvadi *et al.*, 1997). The linker regions of the class I and class II P-glycoproteins, which are capable of conveying drug resistance, each contain consensus sequences for cAMP- and cGMP-dependent protein kinase phosphorylation sites, while

the class III members do not (Hsu *et al.*, 1989). Similar sites are found in the first cytoplasmic domain of the two human P-glycoproteins but not in the murine or hamster homologs.

Recently, the topography of the P-glycoprotein predicted by the hydrophobicity analysis has been questioned (Beja and Bibi, 1995; Cianfriglia *et al.*, 1996; Skach *et al.*, 1993; Zhang *et al.*, 1995). Investigations of the membrane orientation of P-glycoprotein using cell free translation/translocation systems and fusion proteins in bacteria have suggested that multiple topological forms may be present. Polypeptides with four transmembrane domains in either or both the amino half or the carboxy half have been proposed which is distinct from the six transmembrane domains predicted in each half of the protein by the hydropathy analysis. A soluble, heat-labile and trypsin-sensitive activity present in wheat germ extract appeared to modulate different topological structures (Zhang and Ling, 1995). These results are contrasted by P-glycoprotein topology mapping based on the insertion of a hemagglutinin epitope in the predicted intra- or extra-cellular domains (Kast *et al.*, 1996). Immunological localization of the hemagglutinin epitope on functionally expressed P-glycoprotein is consistent only with the twelve transmembrane spanning domain model predicted from the hydropathy analysis.

Insight on the structure of P-glycoprotein has recently been revealed by electron microscopy and single particle images and Fourier projection lamps of 2-dimensional crystalline arrays. The initial interpretation of this data suggests that the P-glycoprotein is a cylinder of approximately 10 nm in diameter with one half of the molecule in the lipid bilayer and the remainder above and below the membrane. There is a large central pore of approximately 5 nm in diameter surrounded by a roughly hexagonal array of the membrane spanning segments. The pore is aqueous and larger than is required for the passage of known substrates. It appears to be closed at the cytoplasmic face by the two intracellular domains which contain nucleotide binding folds. The large size of the pore may explain why the nucleotide binding domains are accessible from the outside of the cell (Baichwal *et al.*, 1993). Attenuated total reflection infrared spectroscopy analysis similarly indicated that a large fraction of the P-glycoprotein is poorly accessible to the aqueous phase and that MgATP, but not ligands, induced tertiary structural changes which increased accessibility to this compartment (Sonveaux *et al.*, 1996).

Mdr5'	M..E.....F.....K.K.....FRY..W.D.L.M..GT..A..HG..LPL.M..FG.MTD.F.....	100
Mdr3'	R..S.....P..SF.....LN..EWPY..VG..CA..NG..QP.....S.....F.....	764
	LEE.M..YAYYY.G.G.GVL..AYIQVS.W.LAAGRQI.KIR..FFHAI..QE.GWFD.....E..LNTRLTDD.SKI..GIGDK..	200
	N.FSL.FL..G...F.T.F.QGFTFGKAGEILT.R.R.M.F..MLRQD.SWFD...N.TG.L.TRLA.D...V.GA....L	864
	GMFFQ...TF...FI.G..F..GWKLTLLVI.A.SP..GLS...WAK..L..F..KEL.AYAKAGAVAE...AIRTVIAFGGQ.KELERY...LE.AK..G	300
	A...QN.ANLGTG.I.S....YQWQLTLLL...P.I...G..EMK.L.G.A..DKK..E..GKIATEAIEN.RT.VSLT.E.KFE.MY...L..PYRN.	964
	IKKA..A.IS.G.A.LL.YASYALAFWYG..LV.S.EY..G...TVFF..L.G.FS.G...P.I..FANARGAA...F.IIDN.P.IDS.S..G.KPD..	400
	...AH..GITF...QA..YFSYA.CFRFG.YL.....M.F..V.LVFS.A.V.GA.A.G..SSFAPDYAKA..SA.....E..P.IDSYS..GL.P...	1064
	.GNL.F....F.YPSR....ILKGLNLKV.SGQTVLVG.SGCGK.TT.QL.QRLYDP..G....DGQDIR..NV..LREIIGVVSQEPVLF.TTIAENI	500
	EG...F....FNYPTR...PVLQGL..EVKKGQTL.LVGSSGCGKSTVVQLLEFY.P.AG.V.LDG.E...LN.Q..RA.LGIYSQEP.LFDCSIAENI	1164
	.YG..R..VTM.EI.KAVKEANAY.FIMKLP..FDTLVG.RGAQLSGGQKQRIARALVRNPKILLDEATSALDTESEA.VQ.ALDKAR.GRTTIVIA	600
	AYGDNRS.V...EI..AA..ANIH.FI..LP.KY.T.VGDKGTQLSGGQKQR.AI.RAL.RQP..LLLDEATSALDTESEKVVQEAQKAREGRTC.VIA	1264
	HRLST.RNAD.IAGF..GVIVE.G.H.ELM...G.Y..LV..QT.G.....	651
	HRLSTIQNADLIVV..NG.VKEHGTHQQLLAQKGIYFSM...QAG.....	1315

Figure 2. Consensus sequence for mammalian P-glycoproteins. Mammalian P-glycoproteins listed in Table III were aligned using the UWGCC Pileup program (gap weight of 3.0 and a gap length weight of 0.1) and the identical amino acids listed using the single letter code. The proximal and distal halves were aligned using the Gap program (gap weight 5 and gap length weight 0.3). Dots represent non-identical residues or gap. Transmembrane domains identified under the single lines and the nucleotide binding domains under the double lines (Endicott, *et al.*, 1991). Shaded amino acids represent identity consensus of the P-glycoproteins described in the text excluding PFMDR1, STE6, P-glycoprotein-like and miscellaneous homologs.

Although not supported by the electron microscopy, a growing body of evidence suggests that the P-glycoprotein may function as a dimer or oligomer. Freeze fracture analysis of multidrug resistant cell lines demonstrates an increase in the density of intramembrane particles with a diameter predicting a molecular mass of 340 kD (Sehested *et al.*, 1989). The particles cluster in the presence of verapamil, although not doxorubicin, suggesting they may represent localization of the P-glycoprotein. These results are complemented by radiation inactivation analysis of membrane preparations from multidrug resistance cell lines indicating that the P-glycoprotein has a molecular mass of 250 kD, about twice the size of the unglycosylated polypeptide (Boscoboinik *et al.*, 1990). Finally, immunoprecipitation of P-glycoprotein from multidrug resistant cells incubated with a cross-linking agent identifies a minor band at 340 kD in addition to the major 180 kD species (Naito and Tsuruo, 1992). The 340 kD band is photolabelled with ³H-Azidopine, but only the 180 kD species is detected when the cells are treated with a reversible cross linking reagent. Finally, thermodynamic characterization of purified P-glycoprotein in defined lipid bilayers similarly suggest aggregation or oligomerization (Romsick and Sharom, 1997). These observations provide indirect evidence for a quaternary structure of the P-glycoprotein as a dimer in the membrane.

Homology amongst the mammalian P-glycoproteins

The mammalian P-glycoproteins display a high level of primary structure similarity with 69–94% amino acid identity and up to an additional 10% conserved substitutions (Figures 4A). The highest similarities exist within each of the P-glycoprotein classes (87–94% identical). The family members which convey multidrug resistance are more similar to each other than they are to the homologs which do not convey resistance both within and across each species. A lower than expected silent nucleotide replacement has suggested that hamster class I and II may not have evolved independently from each other and that gene conversion may have occurred to account for high levels of nucleic acid homology within each of the species (Endicott *et al.*, 1991; Hsu *et al.*, 1989; van der Bliek *et al.*, 1988).

A graphical representation of the amino acid similarity among the P-glycoproteins can be displayed using the Plotsimilarity analysis (Devereux *et al.*, 1984). In this analysis, amino acids are compared at each position in a series of aligned polypeptides. A score is given for the amount of similarity within a window of 10 residues and plotted along the length of the polypeptide. A value of 1.5 is given for identical amino acids, while lesser values are given for homologies based on the Dayhoff amino acid substitution table. Plotsimilarity analysis of all mammalian

Figure 3. Amino acid identity between the amino and carboxy half of the P-glycoprotein homologs. The two halves of the P-glycoproteins aligned using the UWGCG Gap program. Abbreviations for described homologs: H, human; CH, Chinese Hamster, M, mouse, E, *E. histolytica*, Ar, *atpgp1*, Ce, *C. elegans*, Sm, *S. mansoni*, D, *Drosophila*, Pmd, *pmd1*⁺, Le, *lemdr1*, Ld, *ldmdr1*, Pf, *pfmdr1*, St, *STE6*, Lt, *ltgpa*, CF, *CFTR*, Pd5, *PDR5*, Snq, *SNQ2*, Sur, *SUR*, Yor, *YOR1*, Bfr, *BFR1*, Ycf, *YCF1*.

Figure 4. Amino acid identity between each of the P-glycoprotein homologs and (A) human *Mdr1* or (B) *Drosophila Mdr49*. P-glycoprotein homologs in Table III were aligned using the UWGCG Pileup program and identical residues determined using the Distances program. Abbreviations as in Figure 3.

P-glycoproteins (Figure 5A) indicates that the highest levels of similarity are located in the cytoplasmic domains (amino acids 400–600 and 1100–1300). Several notable regions account for the majority of the difference between the family members. The region with the largest amount of amino acid divergence is located in the first external domain, in the region of amino acids 80–120. A similar, but not quite as extensive, divergence is also present in the corresponding region in the carboxy half of the polypeptide, external domain four, encompassed by amino acids 720–760. Finally, the cytoplasmic domain at the amino terminus and the corresponding region at the beginning of the carboxy half of the polypeptide (amino acids 640–700) are quite divergent among the P-glycoproteins.

Plotsimilarity analyses of P-glycoprotein class I (Figure 5B), class II (Figure 5C) and class III (Figure 5D) indicates a similar pattern of amino acid similarity along the polypeptide. As expected, based on higher homologies within each class of P-glycoprotein, amino acid homology along the entire length of the polypeptide is higher with values closer to 1.5, the value for identity. In addition, the regions of highest similarity and divergence are located in similar regions found in the comparison of all the mammalian homologs. However, several notable features are observed. The degree of divergence in the first external domain present in class I and class II P-glycoproteins is almost as high as found across all of the classes (Figure 5A). In this region, around amino acids 80–120, the general increased homology within the class does not appear to apply. This divergence is found between the murine and hamster as well as the human homologs (data not shown). On the other hand, there is very little divergence in the amino acid sequence of the first or fourth external domain in the class III P-glycoproteins (Figure 5D, amino acid regions 80–120 and 720–760). This high level of amino acid conservation in the first external domain of the human, mouse and hamster class III P-glycoproteins is quite remarkable since this is the region of the polypeptide which displays the highest level of divergence amongst the class I and II P-glycoproteins. This is also the same region which displays the most divergence among the different classes within each species (data not shown). Specific functional requirements appear to have resulted in differential evolution of this domain. Finally, the high level similarity along the entire length of the P-glycoproteins, exemplified by scores approaching 1.5, is indicative that the analogous amino and carboxy halves display considerably higher similarity

than the two halves of any individual P-glycoprotein, both within and across the different classes.

P-glycoprotein expression

The identification of molecular and immunological probes for the different classes of P-glycoprotein provided a means to precisely examine their expression in multidrug resistant cell lines and normal tissues. Most multidrug resistant cell lines generated by stepwise selection overexpress P-glycoprotein at levels proportional to the amount of drug resistance. Characterization of human multidrug resistant cell lines indicates that at relatively low levels of resistance, *mdr1* transcripts are overexpressed prior to DNA amplification (Lemontt *et al.*, 1988; Shen *et al.*, 1986). Induction of P-glycoprotein has been demonstrated with a variety of noxious stimuli suggesting that this may be a generalized response to cellular stress (Chaudhary and Roninson, 1993). As resistance increases with selection to higher drug concentrations, an increased number of *mdr1* gene copies are present. RNA overexpression, however, remains increased out of proportion to the amount of gene amplification, often exceeding a 10 fold excess in transcripts (Shen *et al.*, 1986; Van der Bliek *et al.*, 1988). Although the *mdr3* gene is often amplified at levels similar to *mdr1*, RNA expression is usually not detected.

Expression of both class I and II P-glycoproteins correlates with multidrug resistance in murine cell lines (Hsu *et al.*, 1989; Raymond *et al.*, 1990). At relatively low levels of resistance, murine lines which overexpress class I P-glycoprotein have increased *mdr* transcripts prior to gene amplification. Highly resistant lines express more class I RNA than would be expected from the amount of DNA amplification, similar to the human cell lines. In multidrug resistant cell lines which overexpress class II P-glycoprotein, however, both RNA expression and gene amplification are proportional and roughly parallel the level of resistance. The genes for both class I and II P-glycoproteins are consistently amplified in highly resistant cell lines, however, differential levels of transcriptional activation indicate that they are independently regulated. High levels of either or both class I and II P-glycoproteins may be present and the relative contribution may change with the level of selective pressure. The gene encoding class III P-glycoprotein is often amplified to similar levels, however, expression is quite variable and does not correlate with resistance.

Figure 5. Plotsimilarity analysis of mammalian (A), class I (B), II (C), and III (D) P-glycoproteins. Mammalian P-glycoproteins listed in Table III aligned with UWGCG Pileup program and the running average of the amino acid similarities plotted along the length of the polypeptides as described in text. Dashed line represents average similarity. Position in amino acids.

Analysis of RNA (Baas and Borst, 1988; Chin *et al.*, 1989; Croop *et al.*, 1989; Fojo *et al.*, 1987; Trezise *et al.*, 1992) and protein expression (Bradley *et al.*, 1990; Cordon-Cardo *et al.*, 1990; Georges *et al.*, 1990; Sugawara *et al.*, 1988; Thiebaut *et al.*, 1987; Thiebaut *et al.*, 1989) indicates that the P-glycoprotein genes are differentially expressed in normal mammalian tissues. Either or both classes I and II in mice and hamsters are expressed in tissues where human *mdr1* is expressed. P-glycoprotein is consistently expressed at high levels at a variety of secretory surfaces including the bile canaliculi (Buschman *et al.*, 1992), the proximal tubules of the kidney, and intestinal and colonic epithelium. P-glycoprotein is also detected at the epithelial surfaces of the pancreas, bronchial mucosa, ovarian follicles and prostate. A dramatic increase in P-glycoprotein expression occurs in the murine uterus during gestation (Arceci *et al.*, 1990; Arceci *et al.*, 1988). The low levels normally present in the epithelium lining the uterine cavity and secretory glands gradually increase under the combined control of estrogen and progesterone to the highest levels found in mouse tissues. Expression of the cystic fibrosis gene (see below) appears to complement that of *mdr* in a variety of secretory epithelial tissues, including the uterus. The high levels of expression of the cystic fibrosis gene found in the non-pregnant uterine epithelium decrease to very low levels during pregnancy

(Trezise *et al.*, 1992). In addition, CFTR deficient mice express higher levels of P-glycoprotein in the intestine, further suggesting co-ordinate regulation and possible complementary function of the two genes (Trezise *et al.*, 1997).

P-glycoprotein expression is high in other specialized tissues in addition to those with epithelial secretory surfaces. Very high levels of expression are found in the adrenal gland, primarily in the cortical regions. Observations in hamsters indicate that the high level of adrenal expression is limited to males, suggesting P-glycoprotein may be involved in the transport of sex specific hormones (Bradley *et al.*, 1990). P-glycoprotein is also expressed in capillary endothelia in a variety of tissues. Expression is most notable in capillaries found in the central nervous system, testes, uterus, and the papillary dermis of the skin, which has prompted speculation that P-glycoprotein may play a role in establishing the blood brain barrier (Cordon-Cardo *et al.*, 1989). P-glycoprotein expression has also been observed in hematopoietic stem cells (Chaudhary and Roninson, 1991) and mature lymphoid cells (Chaudhary *et al.*, 1992; Drach *et al.*, 1992). Finally, class III P-glycoprotein is specifically expressed in human liver, spleen and kidney as well as in murine adrenal and heart – tissues where classes I and/or II are also expressed.

Analysis of P-glycoprotein in normal tissues in

conjunction with mice lacking normal P-glycoproteins has shed significant insight on its normal physiological role. Characterization of mice with a disrupted *mdr1a*, *mdr1b*, or both *mdr1a* and *mdr1b* have suggested that the class I and II P-glycoproteins protect the organism from exogenous xenobiotics (Mayer *et al.*, 1996; Schinkel *et al.*, 1997; Schinkel *et al.*, 1994; Schinkel *et al.*, 1995; Sparreboom *et al.*, 1997; van Asperen *et al.*, 1996). Physiological abnormalities can not be identified in any of the deficient mice which appear to have normal growth, development and fertility. It is not clear if another protein can complement some or all of the functional activities of the P-glycoprotein (see CFTR below). However, administration of P-glycoprotein substrates results in a two to four fold increase in the concentration of the substrates in the organs of these mice. Most notably there may be a ten to one hundred fold increase in the concentration of the substrate in the brains of the mice lacking *mdr1a*. This is consistent with the observation that high levels of *mdr1a* P-glycoprotein are expressed in the capillary beds of the brain as well as the testes and skin. This has suggested that P-glycoprotein may be involved in the maintenance or generation of the blood brain barrier. In addition, analysis of the oral administration of paclitaxel to mice deficient in *mdr1a* indicated that P-glycoprotein normally limits the oral uptake of substrates in the multidrug resistant phenotype. Furthermore, although biliary excretion is not different between wild type and *mdr1a* deficient mice, significantly less paclitaxel is excreted into the bile suggesting a role for P-glycoprotein in enteric elimination of specific substrates. Similarly, analysis of digoxin pharmacokinetics indicated that the predominant fecal excretion in wild type mice shifted to predominantly urinary excretion in the P-glycoprotein deficient mice.

Mice with a disrupted murine *mdr2*, class III P-glycoprotein, which does not convey multidrug resistance, have abnormal hepatic architecture and decreased phospholipid in the bile (Smit *et al.*, 1993). Analysis of the murine and human class III P-glycoprotein indicate that this homolog functions as a phospholipid translocase (Ruetz and Gros, 1994). Recently human MDR1 has also been shown to function as a phospholipid translocase although with broader specificity than the class III homolog (Bosch *et al.*, 1997; van Helvoort *et al.*, 1996). Interestingly, overexpression of human MDR3 in transgenic mice results in a peripheral neuropathy with slowed motor nerve conduction and dysmyelination of the peripheral nerves

(Smit *et al.*, 1996). In addition, overexpression of human MDR3 in the lens results in a severe lenticular pathology which resemble that identified in congenital mouse osmotic cataract (Dunia *et al.*, 1996). How overexpression of class III P-glycoprotein results in these pathological states is unclear. Recent observations on the ability of the human P-glycoprotein to function as an ATP channel (Abraham *et al.*, 1993; Bosch *et al.*, 1996) and regulator of a volume regulated chloride channel (Hardy *et al.*, 1994) implies multiple functional properties associated with the polypeptide and have challenged the traditional distinctions between transporters and channels (Jan and Jan, 1992).

P-glycoprotein homologs

P-glycoprotein homologs have been isolated from a wide range of species. Analysis of cDNA and genomic clones indicate a high level of amino acid similarity and structural conservation, based on hydrophobicity analyses, present across large evolutionary distances. High levels of similarity even remain in the nucleic acid sequence in many of the distant homologs. Isolation of multiple P-glycoproteins and cross hybridization of genomic DNA to molecular probes indicate the presence of a multigene family in most organisms (Juranka *et al.*, 1989). Although specific functions have not yet been identified for many of the P-glycoprotein homologs, the variability in the number of family members and the divergence in the primary sequence suggest that this transport protein has readily adapted to a range of functional capacities.

The precise requirements to be included as a P-glycoprotein remain elusive. ABC transporters can be comprised of multiple protein subunits yet appear to form a similar transport unit in the membrane (Figure 6). Whether these structural differences underlie significant functional mechanisms or are merely historical relics of convergent or divergent evolution is unknown. Currently, most would include those with high amino acid similarity and similar topology as P-glycoproteins. Figures 3, 4 and 7 show the similarities between the amino acid sequences of the P-glycoprotein homologs and should be referred to for the comparisons described in the text.

Saccharomyces cerevisiae

The yeast *Saccharomyces cerevisiae* *STE6* gene encodes a P-glycoprotein homolog (Kuchler *et al.*, 1989;

Figure 6. Schematic representation of the evolutionary origins and organization of the functional units utilized in P-glycoprotein transport.

McGrath and Varshavsky, 1989). *STE6* is one of several non-lethal 'sterile' mutations in *MATa* cells which inhibit haploid *S. cerevisiae* mating and the formation of diploid cells. *MATa* cells secrete **a** factor, a 12 amino acid peptide pheromone, which induces conjugation. The **a** factor lacks a hydrophobic N-terminal signal sequence and sites for N-linked glycosylation suggesting that it is not exported using traditional secretory pathways. However, **a** factor does undergo post-translational farnesylation and methylation which render it lipophilic. Yeast carrying the *STE6* mutation produce **a** factor but are unable to secrete the peptide to the extracellular space while overexpression of *STE6* results in a striking increase in the amount of **a** factor released from the cells (Kuchler *et al.*, 1989).

Analysis of the predicted structure of the *STE6* protein product reveals a polypeptide with 1290 amino acids comprised of 12 membrane spanning domains and two cytoplasmic domains containing nucleotide binding folds. The hydrophobicity profiles of *STE6* and the mammalian P-glycoproteins are nearly superimposable. The two halves of *STE6* have approximately 27% amino acid identity and a total of 50% amino acid similarity. There is approximately 25% amino acid identity with the mammalian P-glycoproteins along the entire polypeptide, with the highest homology, approximately 40%, in the nucleotide binding domain. The homology to P-glycoprotein has suggested that *STE6* is a transport protein capable of translocating **a** factor to the extracellular space in a fashion distinct from classical secretory pathways. The requirement that **a** factor undergo hydrophobic modification implies that lipophilicity may be a general property of P-glycoprotein transport substrates and that the natural transport products of

the mammalian homologs may undergo similar functional modification (McGrath and Varshavsky, 1989). Preliminary observations suggest that yeast deficient in *STE6* were more sensitive to the peptide antibiotic valinomycin, while overexpression of the gene results in resistance (Kuchler *et al.*, 1989).

Mutations generated in either the proximal or distal nucleotide binding domains of *STE6* inhibit export of **a** factor suggesting that each half of the polypeptide is required for specific functional activities (Berkower and Michaelis, 1991). Similarly, expression of either half of *STE6* is not sufficient for export of **a** factor. Co-expression of the two halves, however, does result in export of **a** factor. Immunoprecipitation indicates that the separate *STE6* moieties are in physical contact (Berkower *et al.*, 1996). These observations have been further extended to indicate that even the N-terminal membrane spanning domains can interact physically and functionally with the remaining three quarter molecule of *STE6*. Interestingly a wild type half molecule can interact and complement a non-functional full length optimal which has a mutation in the corresponding half of the polypeptide. This flexibility in protein interaction supports the concept that all ABC transport proteins have a similar three dimensional structure and that the components can assemble in the membrane.

In spite of significant amino acid divergence between the homologs, murine MDR3, PFMDR1, and human MRP are capable of transporting **a** factor in *STE6* deficient *S. cerevisiae* (Raymond *et al.*, 1992; Ruetz *et al.*, 1996; Volkman *et al.*, 1995). Although export is much less efficient than by *STE6*, functional homology across a large evolutionary distance appears to remain. Similar experiments with human MDR1

indicate that the human homolog is not capable of a factor transport even though homology between the mammalian P-glycoproteins is much higher to each other than to STE6 (Kuchler and Thorner, 1992).

Schizosaccharomyces pombe

A P-glycoprotein homolog has been isolated from a leptomycin B resistant strain of *Schizosaccharomyces pombe* (Nishi *et al.*, 1992). Leptomycin B is an antifungal antibiotic isolated from a strain of *Streptomyces* which induces morphological abnormalities in a variety of fungi including the fission yeast *S. pombe*. A leptomycin B resistant strain of *S. pombe* identified after chemical mutagenesis has been utilized to characterize the genes responsible for resistance. A genomic library from the resistant line was cloned into a multicopy expression vector and introduced into a susceptible stain. Three distinct clones were identified which were capable of conveying leptomycin B resistance. Genetic analysis of the clone which conveyed the highest level of resistance suggested that it was not a mutant gene, but rather a wild-type gene able to confer resistance when carried on a multicopy plasmid. Overexpression of this plasmid resulted in resistance to leptomycin B, cycloheximide, valinomycin and staurosporine B but not actinomycin D. Disruption of the wild-type gene indicated that it was non-essential, however, the null mutant was hypersensitive to the agents to which it conveyed resistance as well as actinomycin D. Analysis of the clone indicated that an open reading frame consisting of 1362 amino acids with similarity to the predicted amino acid sequence of the P-glycoprotein was responsible for resistance. The gene was designated *pmd1*⁺ for *S. pombe* *mdr1*-like gene. The primary structure predicted a polypeptide with two similar halves, each containing six transmembrane domains and a nucleotide binding domain. *PMD1*⁺ has higher amino acid identity to the mammalian P-glycoproteins (37–40%) than to the yeast homolog STE6 (24%).

Plasmodium faciparum

The rapid emergence of *Plasmodium faciparum* resistant to antifolate antimalarials has led to a dependence upon quinoline-containing compounds for the treatment and prophylaxis of malaria. Parasites resistant to these newer agents, however, are increasingly prevalent in most endemic areas of the world, posing a serious threat to control of the disease. Analysis of

resistant *P. faciparum* initially suggested that the organisms displayed a phenotype similar to mammalian multidrug resistant cell lines. Isolates found in the wild were shown to be resistant to multiple antimalarials such as chloroquine, mefloquine and quinine (Martin *et al.*, 1987). In addition, *P. faciparum* isolates resistant to chloroquine were shown to efflux radiolabelled chloroquine 40–50 time more rapidly than susceptible stains, even though the initial uptake rates were similar (Krogstad *et al.*, 1987). Finally, verapamil and diltiazem, reversal agents of mammalian multidrug resistance, as well as vinblastine and doxorubicin all were shown to increase the accumulation of chloroquine in the resistant but not the susceptible isolates (Krogstad *et al.*, 1987; Martin *et al.*, 1987). These observations suggested that a P-glycoprotein homolog may have been involved in the resistance phenotype of *P. faciparum*.

Evidence for P-glycoprotein homologs in *P. faciparum* was obtained by screening genomic libraries using degenerate oligonucleotide probes to sequences in the highly conserved nucleotide binding fold (Foote *et al.*, 1989) and analysis of polymerase chain reaction products which were designed to span the nucleotide binding fold (Wilson *et al.*, 1989). The complete coding region of *pfmdr1* reveals a polypeptide of 1420 amino acids in length consisting of two homologous halves that each contain six transmembrane regions and a nucleotide binding domain (Foote *et al.*, 1989). A stretch of approximately 50 additional amino acids, not present in other P-glycoproteins, is inserted between the two consensus sequences which comprise the nucleotide binding folds. The entire polypeptide displays approximately 30% amino acid identity to the mammalian homologs and each of the two homologous halves are more similar to the mammalian homologs than they are to each other.

Characterization of the role of *pfmdr1* in antimalarial resistance has been hampered by the lack of a genetic model in *P. faciparum*. The recent technological advance of transfecting *P. faciparum* should hopefully extend the analysis of drug resistance in this parasite (Wu *et al.*, 1995). Conflicting evidence has confounded its role in drug resistant *P. faciparum*. Over half of the analyzed chloroquine resistant isolates have amplified *pfmdr1* (Price *et al.*, 1997; Triglia *et al.*, 1991; Wilson *et al.*, 1993). Quantitative immunoblotting indicates that the protein is not overexpressed in resistant isolates unless the gene has been amplified (Cowman *et al.*, 1991). Direct evidence for the role of *pfmdr1* in drug resistance was demonstrated

in the expression of PFMDR1 in yeast cells (Ruetz *et al.*, 1996). Yeast expressing PFMDR1 displayed resistance to quinacrine, mefloquine, and halofantrine. Resistance was associated with decreased accumulation of the anti-malarial agent quinacrine. Introduction of amino acid polymorphisms associated with drug resistance abolished the capacity of the protein to confer drug resistance. These observations, along with the overexpression of PFMDR1 in mammalian cells conveying sensitivity to chloroquine, support the hypothesis that PFMDR1 may be transporting the anti-malarial agents into an intracellular compartment such as the digestive vacuole for removal of the toxic compounds.

Analysis of the *pfmdr1* gene in chloroquine resistant isolates indicated the presence of two alleles correlating with resistance (Foote *et al.*, 1990). Amino acid substitutions occur at locations where independent mutations affect the ability of the mammalian P-glycoprotein to transport specific cytotoxic substrates out of the cell (Choi *et al.*, 1988; Gros *et al.*, 1991). Examination of the predicted amino acids at these positions correctly identified whether 34 of 36 *P. falciparum* isolates were resistant. Analysis of patients treated with either chloroquine or amodiaquine who remained positive for the parasite seven days after treatment indicated that there is a significantly higher prevalence of the TYR-86 allele in the parasites suggesting that this allele conveys resistance to both of these anti-malarials (Duraisingh *et al.*, 1997).

The original characterization of a chloroquine resistant isolate, W2, which had subsequently been placed under stepwise selection of mefloquine *in vitro*, W2mef, indicated that amplification and overexpression of *pfmdr1* was associated with mefloquine resistance but not chloroquine resistance (Wilson *et al.*, 1989). Evaluation of *P. falciparum* from patients failing standard therapies has similarly suggested a correlation between mefloquine resistance and *pfmdr1* amplification (Wilson *et al.*, 1993). Although there was no correlation between the presence of the alleles described above and chloroquine resistance, amplification of the *pfmdr1* gene was found in 10 of the 11 isolates. All 10 of the isolates which had amplified *pfmdr1* were resistant to all compounds tested. The only isolate which did not display amplification of *pfmdr1*, however, was sensitive to mefloquine and halofantrine.

A role for *pfmdr1* in conveying resistance to anti-malarials in *P. falciparum* could not be found utilizing a genetic cross between chloroquine resistant and sen-

sitive parasites (Wellems *et al.*, 1990). The cross was made between a chloroquine sensitive line, HB3, containing a single copy of *pfmdr1*, and a resistant line, Dd2, a derivative of W2mef, which contains 4 copies. Eight chloroquine resistant and eight sensitive isolates were identified and characterized *in vitro*. Each of the progeny clones displayed a sensitive or resistant phenotype identical to one of the parents. Neither the number of copies of Dd2 *pfmdr1*, or even the presence of Dd2 *pfmdr1* correlated with chloroquine resistance. Resistant and sensitive isolates were just as likely to receive *pfmdr1* from either parent. Similarly, there were no significant differences among the parent or progeny clones in the level of mefloquine resistance. Analysis of the sensitive HB3 parent, however, has indicated that it has one of the alleles associated with chloroquine resistance and in addition, may be somewhat resistant to mefloquine (Foote *et al.*, 1990). The analysis of this cross could be complicated if the HB3 *pfmdr1* gene has the ability to convey resistance. Although this cross does not appear to implicate *pfmdr1*, it does suggest, however, that a single genetic locus is responsible for chloroquine resistance.

Leishmania

Selection of *L. donovani* with increasing concentrations of vinblastine have successfully generated parasite lines cross resistant to vinblastine, daunorubicin and puromycin (Henderson *et al.*, 1992). The resistant parasites display markedly decreased uptake of radiolabelled puromycin. Drug resistance, however, was not reversed with either verapamil or quinidine even though the parasites displayed a multidrug resistant phenotype similar to mammalian multidrug resistant cells. A 5.7 kbp genomic clone from *L. donovani* with the complete coding sequence of *ldmdr1* was isolated using a probe derived from a polymerase chain reaction product generated to identify a nucleotide binding domain. The clone was transfected into *Leishmania* and amplified by co-selection with the neomycin resistance gene. The transfected parasites displayed the same multidrug resistant phenotype as the original resistant organisms. An analogous gene, *lemdr1* has been isolated from *L. enriettii* which is amplified on a 35–40 kbp extrachromosomal circle in *Leishmania* selected in vinblastine (Chow *et al.*, 1993). Analysis of the predicted amino acid sequences reveal polypeptides with a hydrophobicity profile similar to the mammalian P-glycoproteins, 83% amino acid identity with

each other and approximately 37% amino acid identity to the mammalian homologs.

Drosophila

Three members of the *Drosophila mdr* multigene family have been isolated and characterized. The P-glycoprotein homologs have been designated relative to their chromosomal location, *Mdr49* on chromosome two, *Mdr65* on chromosome three (Wu *et al.*, 1991), and *Mdr50* also on chromosome two, but far from *Mdr49* (Gerrard *et al.*, 1993). This is in contrast to the mammalian multigene family which is clustered on a single chromosome. A fourth *Drosophila* homolog is located in tandem with *Mdr65*, and has been designated *Mdr65b* (J. Croop, unpublished observations). The predicted amino acid sequence for all three P-glycoprotein homologs indicate that the polypeptides have two homologous halves, each with six transmembrane domains and a nucleotide binding fold. The *Drosophila* homologs display approximately 40–45% amino acid identity to the mammalian homologs. However, the nucleic acid and amino acid sequences of the *Drosophila* P-glycoproteins which are on different chromosomes are as divergent from each other as they are from the mammalian homologs, in contrast to the mammalian *mdr* families which display a high level of intraspecies similarity (compare Figures 4A and B). MDR65 and MDR65b have the highest similarity amongst the *Drosophila* homologs. Functional analysis of both MDR49 and MDR65 indicate that both *Drosophila* P-glycoprotein homologs can transport P-glycoprotein substrates and bind traditional P-glycoprotein photoaffinity probes (I. Bosch and J. Croop, manuscript in preparation). Preliminary observations indicate that *Mdr49* and *Mdr65* are differentially expressed at the secretory surfaces of tissues analogous to mammalian structures where P-glycoprotein is expressed (R. Arceci and J. Croop, unpublished data).

Mdr49 is located in a chromosomal region which has been the subject of previous genetic analysis (Wu *et al.*, 1991). Three deficiencies were identified which had chromosomal breakpoints within the *Mdr49* gene. *vg*¹³⁵ had a breakpoint in the proximal coding region, *vg*^C near the center of the gene and *vg*¹³⁸ probably just distal to the coding region. In addition, the *b8* line deleted the entire gene. Each of the deficiencies is a large chromosomal deletion which is homozygous lethal. However, genetic crosses in flies with deleted and disrupted *Mdr49* resulted in progeny lacking a

wild-type *Mdr49* gene. Utilizing the balancer chromosome, *CyO*, which gives flies curly wings, *vg*¹³⁵/*CyO* and *vg*^C/*CyO* flies were mated with *b8*/*CyO* flies. The flies with straight wings would be the *vg*¹³⁵/*b8* and *vg*^C/*b8* progeny with *Mdr49* deleted on the *b8* chromosome and disrupted on the other. Molecular analysis indicated loss of chromosomal material and truncated mRNA species consistent with the predicted genetic localization.

Functional analysis of the mutant *Mdr49* fly lines indicated that they were fertile and viable under laboratory conditions. The effect of losing the *Mdr49* gene in flies stressed with a cytotoxic agent was ascertained by determining the survival of flies developing on colchicine. The number of adult flies emerging was determined for 19 randomly chosen control lines and the flies bearing a disrupted *Mdr49*. The *vg*^C/*b8* flies were approximately two fold more sensitive to colchicine than the controls and could easily be distinguished from survival of any of the other lines at 10uM colchicine. Neither the *vg*¹³⁵/*b8* or the heterozygous *vg*¹³⁵, and *vg*^C, and *b8* lines showed an increased sensitivity to colchicine. Thus the loss of the wild-type *Mdr49* in and of itself did not cause the heightened colchicine sensitivity during development. It is possible, however, that the remaining portion of the disrupted *Mdr49* gene gives rise to a truncated or fusion protein or RNA species which is deleterious to development on colchicine. The *vg*^C RNA transcript from the flies with increased sensitivity includes almost half of the *Mdr49* coding sequence while the *vg*¹³⁵ transcript is very short and unlikely to provide deleterious information.

Entameoba histolytica

Entameoba histolytica is an enteric protozoan responsible for amoebic dysentery in man. Although drug resistance has not been a major therapeutic problem, case reports of failed drug treatment have been described. *E. histolytica* mutated with ethylmethanesulphonate and selected in emetine have resulted in a resistant clone with a stable phenotype. Probes generated from polymerase chain reaction products predicted to span the nucleotide binding fold identified four distinct DNA fragments containing P-glycoprotein-like sequences (Samuelson *et al.*, 1990). The probes identified 4.5–5 kbp mRNAs overexpressed in the emetine resistant line. When the resistant line was selected with increasing concentrations of emetine, higher levels of mRNA expression were

observed. The sequences for two P-glycoprotein homologs, *Enhpgp1* and *Enhpgp2*, have been characterized in detail (Descoteaux *et al.*, 1992). The two homologs are 1302 and 1310 amino acids in length displaying a higher level of amino acid identity to each other, 67%, than to other P-glycoprotein homologs. This is in contrast to *Drosophila*, *Leishmania*, and *C. elegans* where similarities between P-glycoproteins within the same organism are no greater than the similarities to homologs in other species. Each of the *E. histolytica* homologs display approximately 38–40% amino acid identity with the mammalian P-glycoproteins. The predicted structure is very similar to the mammalian P-glycoprotein except that the amino terminus is not as hydrophilic. One of the homologs has been shown to convey emetine resistance in *E. histolytica* after transfection (Ghosh *et al.*, 1996).

Arabidopsis thaliana

A P-glycoprotein homolog has been identified in the plant *A. thaliana* (Dudler and Hertig, 1992). Degenerate oligonucleotides corresponding to highly conserved amino acid sequences in the nucleotide binding fold were used to screen a genomic library. Two distinct classes of genomic sequences were obtained, and one, *atpgp1*, further characterized. The predicted polypeptide was 1286 amino acids in length consisting of two similar halves, each comprised six transmembrane domains and a cytoplasmic domain with the consensus sequences for a nucleotide binding domain. The two halves displayed 37% amino acid identity with each other and 36–39% amino acid identity with the mammalian P-glycoproteins. Northern blot analysis indicated that *atpgp1* was expressed in all parts of the plant examined, including the leaves, roots, flower buds, and open flowers. Highest levels of expression of the 4.5 kbp transcript were found in the peduncles.

Caenorhabditis elegans

Four homologs have been identified in the *C. elegans* *pgp* multigene family (Lincke *et al.*, 1992). Three sets of genomic clones, which each mapped to a different chromosome, were isolated using a human *mdr1* probe. Further analysis of the clones indicated that four P-glycoprotein gene homologs were present in the three genetic loci. cDNAs for two of the homologs, *Celpgpa* and *Celpgpc*, have been characterized in detail. They encode polypeptides 1321 and 1254 amino acids in length. Each is comprised of two homologous halves containing six transmembrane domains

and a cytoplasmic region with a nucleotide binding fold. The hydrophobicity plots of both homologs are virtually superimposable on that of human *mdr1*. The two homologs, however, have only 40% amino acid identity with each other and 38–40% amino acid identity with the mammalian P-glycoproteins. The two genes have 14 and 13 exons, respectively. Although the *C. elegans* homologs share only one intron/exon junction, there are four and five intron positions, respectively, conserved with the mammalian *mdr* genes. Partially processed RNA transcripts for one of the P-glycoprotein homologs increased with growth of the nematodes at heat shock temperatures over an eight hour period. There were no increases with arsenite or actinomycin D, however, prolonged exposure to emetine appears to have resulted in increased expression. Disruption of one of the two homologs expressed in the intestine increases sensitivity of the worms to colchicine and chloroquine suggesting a role in protection from environment toxins (Broeks *et al.*, 1995; Lincke *et al.*, 1993).

Xenopus laevis

A P-glycoprotein homolog has been isolated from *Xenopus laevis* by homology to the murine cDNA as a probe (Castillo *et al.*, 1995). The predicted protein sequence is 68 percent identical to the human MDR1 with a similar hydrophobicity analysis. The RNA is expressed predominantly in the intestine with highest levels in the ileum. Vesicles prepared from the brush border membrane of the intestine efflux cytotoxic compound vinblastine. The efflux was ATP dependent and was decreased by the P-glycoprotein modulator verapamil. These observations suggest that this P-glycoprotein has similar functional activities to the mammalian homolog.

P-glycoprotein-like transporters

Two new groupings of ABC transporters with similarity to the P-glycoprotein have recently been identified. Both groups include all of the known components required for a ABC transporter on a single polypeptide. One group includes CFTR and MRP which have a similar topology with the P-glycoprotein but displayed at best 50% identical and conserved amino acids. A second group includes a family of yeast homologs, PDR5, SNQ2, and BFR1 which are mirror images of the P-glycoprotein. Each half of the polypeptide is comprised of a nucleotide binding domain followed

by the transmembrane domains, the reverse found in P-glycoprotein.

Analysis of the gene responsible for cystic fibrosis indicates that it has amino acid homology and a similar structure to the P-glycoprotein (Riordan *et al.*, 1989). The amino acid sequence of the cystic fibrosis transmembrane conductance regulator (CFTR) predicts a polypeptide 1480 amino acids long with two repeat motifs, each of which consists of six membrane spanning domains and a hydrophilic region containing the consensus sequences for a nucleotide binding fold. The two halves of the polypeptide display only 21% amino acid identity with each other and 15% amino acid identity with the mammalian P-glycoproteins. The only homology surrounds the nucleotide binding domain. A large, highly charged cytoplasmic domain of 241 amino acids, termed the R group, contains a cluster of consensus sequences for protein kinase A and protein kinase C phosphorylation sites and links the two halves of the molecule at the analogous position of the 'linker' region in the P-glycoprotein. Although CFTR is best characterized as a chloride channel, overexpression has been demonstrated to convey resistance to cytotoxic drugs included in the multidrug resistance phenotype (Wei *et al.*, 1995). P-glycoprotein and CFTR expression appears to be regulated in a co-ordinated fashion. P-glycoprotein expression complements CFTR expression in a number of epithelial tissues and is increased in the intestine of 3–4 week old CFTR deficient mice (Trezise *et al.*, 1997; Trezise *et al.*, 1992).

The multidrug resistance-associated protein, MRP, has been identified as the mediator of a cross resistant phenotype in cells which have been isolated by stepwise selection in doxorubicin which is distinct from that of P-glycoprotein overexpression (Cole *et al.*, 1992). Molecular analyses indicate that the gene encoding MRP is amplified and overexpressed in the resistant cell line. Although the line displays resistance to multiple agents, traditional P-glycoprotein reversal agents do not circumvent resistance and major differences in net drug accumulation do not appear to be part of the resistant phenotype in selected cell lines. Analysis of the predicted amino acid sequence of MRP indicates a low level of sequence similarity to the P-glycoprotein with two similar halves each containing a nucleotide binding domain, the areas of highest similarity. The hydrophobicity analysis suggests 12 transmembrane domains in the proximal and six in the distal half of the polypeptide (Lautier *et al.*, 1996). A cluster of potential transmembrane domains

at the amino terminus is distinct from all of the other P-glycoprotein homologs. Although the proximal and distal halves are similar, the homology is primarily located around the nucleotide binding domain and the two halves are not as similar to each other as most P-glycoprotein homologs (Figure 3).

Consistent with the differences in primary structure is the suggestion that MRP transports drugs by a different mechanism than P-glycoprotein. Drug transport appears to occur in the presence of reduced glutathione and, in addition, MRP has been demonstrated to have a role in the export of certain glutathione conjugates and cysteinyl leukotrienes (Lautier *et al.*, 1996). Many of the substrates are anionic in contrast to the cationic substrates of the P-glycoprotein (Tusnady *et al.*, 1997). However, like P-glycoprotein, MRP has been shown to be the member of a multigene family (Kool *et al.*, 1997). Decreased expression of a liver specific MRP homolog, the canalicular multi-specific organic anion transporter, cMOAT, has been associated with a specific defect in excretion of conjugated bilirubin and organic acids in rats (Paulusma *et al.*, 1996). Mutations in the gene may be responsible for the unconjugated hyperbilirubinemia found in the human Dubin-Johnson syndrome (Paulusma *et al.*, 1997).

Leishmania is an intracellular parasite of macrophages responsible for a range of clinical syndromes including cutaneous, mucosal and visceral lesions. The treatment of choice remains pentavalent antimonials. One mechanism of resistance in *Leishmania* selected with heavy metals, methotrexate, primaquine, or terbinafine correlates with amplification of H circles (Ellenberger and Beverley, 1989; Ouellette *et al.*, 1990). H circles are extrachromosomal genetic elements comprised of a 30 kbp inverted repeat separated by unique DNA segments which encode at least 20 polyadenylated RNAs. Amplification occurs from pre-existing circles or the *de novo* generation of H circles from a chromosomal copy (Ouellette *et al.*, 1991). *ltgpa* was first identified as a P-glycoprotein homolog in amplified H-circles from *L. tarentolae* using cross hybridization to a mammalian *mdr1* probe (Ouellette and Borst, 1995). Sequence analysis of *ltgpa* indicates a continuous open reading frame of 1548 amino acids encoding a polypeptide with two similar halves, each containing six transmembrane domains and a nucleotide binding fold. Similarity to mammalian P-glycoproteins, however, extends only to the 200 amino acids surrounding the nucleotide binding domain as indicated by only 15% amino acid

identity with the entire polypeptide. There does not appear to be significant similarity between the two halves of LTPGPA outside of the nucleotide binding domain with only 21% amino acid identity between the halves. In addition, the amino terminus is much longer than the corresponding region in the mammalian homologs. These differences, as well as closer homology to other P-glycoprotein-like transport proteins such as MRP have suggested that LTPGPA may be distinct from the other P-glycoproteins (Cole *et al.*, 1992).

Isolation and expression of four DNA fragments spanning the H region of *L. major* have implicated *lmpgpA*, the homolog of *ltgpA*, as conveying resistance to the heavy metals arsenite, antimony and lead (Callahan and Beverley, 1991). Introduction of the H region fragment which includes *lmpgpA* and further amplification by co-selection with the neomycin resistance gene confers heavy metal resistance to the parasite. There was no increase in resistance to cytotoxic agents included in the mammalian multidrug resistant phenotype such as vinblastine and puromycin. Resistance was associated with a two fold decrease in arsenite accumulation. Similarly, when clones which included *ltgpA* were introduced into the parasites, a similar pattern of heavy metal resistance was obtained although the level was approximately 2–4 fold less than with the *lmpgpA* constructs. Similarly, the ABC transporter HMT1 has recently been shown to convey resistance to heavy metals (Ortiz *et al.*, 1992). Analysis of an H circle in *L. tarentolae* amplified during selection with arsenite revealed a novel fusion of *ltgpA* and a second P-glycoprotein homolog designated *ltgpB* which resides just outside the chromosomal location of the H region (Ouellette *et al.*, 1991).

Additional homologs with amino acid similarity to MRP and CFTR have been identified. SUR, the sulfonylurea receptor is an ATP dependent potassium channel which modulates a decrease in serum glucose by stimulation of insulin secretion when its activity is inhibited by oral hypoglycemic agents. Mutations which decrease SUR functional activity have been identified as the etiology of familial persistent hyperinsulinemic hypoglycemia of infancy (Thomas *et al.*, 1995). YCF1, the yeast cadmium resistance factor protects against the toxicity of both cadmium and the glutathione S-conjugate substrate, 1-chloro-2,4-dinitrobenzene (Li *et al.*, 1996; Szczypka *et al.*, 1994). This MgATP dependent vacuolar ATPase is sensitive to both verapamil and vinblastine. MRP functionally complements YCF1 (Tommasini *et al.*, 1996). YOR1,

a yeast ABC transporter with similarity to MRP functions as a multispecific organic anion transporter conveying oligomycin and reveromycin A resistance and overlapping specificity with YCF1 in the resistance to cadmium (Cui *et al.*, 1996; Katzmann *et al.*, 1995).

The family of ABC transporters which include MRP, cMOAT, *LtgpA*, CFTR, SUR, and YOR1 have 25–50% amino acid identity amongst themselves with only about 15% amino acid identity to any of the P-glycoproteins. These observations have suggested that this group is more closely related to each other than the other P-glycoprotein homologs and represent a separate family (Cole *et al.*, 1992). Of note, however, the nucleic acid sequence surrounding the nucleotide binding domain of MRP is more similar to the P-glycoproteins than *LtgpA* or CFTR (data not shown).

A second group of P-glycoprotein-like ABC transporters similarly encode all of the domains required for a functional unit on a single polypeptide. The hydrophilic domain which includes the nucleotide binding fold is proximal to the hydrophobic domain, however, and this configuration is then repeated on the second half of the molecule. Several transporters which export specific or multiple toxins from yeast have been identified (Balzi and Goffeau, 1995). Overexpression of PDR5/STS1/YDR1, SNQ2, and BFR1 convey resistance to multiple antiseptics and other toxic compounds. Each transporter has a distinct drug resistant phenotype, although there is some overlap in specificity. Disruption of these genes results in increased sensitivity to the same compounds. Transcription factors which modulate expression of these multidrug resistance genes have been identified, and interestingly can also modulate transcription of YOR1 which is a MRP homolog (Katzmann *et al.*, 1995). The clinical importance of these homologs has recently been suggested by the transport of steroids by PDR5 (Mahe *et al.*, 1996) and resistance to anti-fungals by CDR1 in yeast (Sanglard *et al.*, 1995).

The predicted structure of the P-glycoprotein is similar to a variety of channels and transport proteins but lacks substantial similarity to their amino acid sequences. A consensus structure consisting of 12 α -helical, transmembrane spanning domains has been proposed for a wide range of transport systems which require energy coupling (Malone, 1990). Primary sequence comparisons have identified the Major Facilitator Superfamily comprised of drug resistance proteins, sugar facilitators, facilitators for Krebs cycle intermediates, phosphate ester-phosphate an-

tiporters and oligosaccharide-H⁺ symporters (Marger and Saier, 1993). Several members of the major facilitator superfamily convey resistance to cytotoxic compounds (Alarco *et al.*, 1997; Edgar and Bibi, 1997; Nourani *et al.*, 1997; Sanglard *et al.*, 1996; White, 1997). In addition, this generalized structure is found in adenylyl cyclase which is comprised of two similar halves each containing six transmembrane domains and a large hydrophilic region with a cyclic nucleotide binding domain (Kupinski *et al.*, 1989). Voltage gated sodium (Noda *et al.*, 1986) and calcium channels (Tanabe *et al.*, 1987) are comprised of four similar groups each including six transmembrane spanning domains. These hydrophobic regions are interspersed with cytoplasmic domains predicting a structure not dissimilar to a pair of linked P-glycoproteins. The P-glycoprotein, however, does not display significant homology at the amino acid level to these transport proteins and channels. Whether the lack of amino acid homology with structural homology represents divergence from the ancestral origins of the transport proteins or convergence to motifs which perform similar functional activities is unknown.

Miscellaneous P-glycoprotein homologs

A P-glycoprotein homolog has been isolated from *Escherichia coli* and designated *Mdl* (Allikmets *et al.*, 1993). It has a similar structure to the P-glycoproteins except that the amino end appears truncated with only five transmembrane domains. There is approximately 20–25% amino acid identity with the other P-glycoprotein homologs. The pig P-glycoprotein multi-gene family includes 5 members (Childs and Ling, 1996). Four of the homologs appear to be related by sequential duplication having highest homology with class I and the remaining homolog having homology to class III P-glycoprotein. A homolog classified as Sister of P-glycoprotein (spgp) has been identified which is primarily expressed in the liver (Childs *et al.*, 1995). Partial sequences for two P-glycoproteins have been isolated from winter flounder using hamster *pgp* probes and have highlighted the difficulty in relating specific homologs to the mammalian classes (Chan *et al.*, 1992). Finally, a P-glycoprotein homolog, *Smdr2*, has been isolated from the parasite *Schistosoma mansoni* using a polymerase chain reaction product generated to span a nucleotide binding domain (Bosch *et al.*, 1994).

Evolutionary homologies

A comparison of the amino acid sequence of the P-glycoprotein to the protein sequences in the genetic databases indicates homology to a rapidly expanding group of polypeptides in both prokaryotes and eukaryotes. The region of homology is highest in the approximately 200 amino acids which surround the nucleotide binding domains. The vast majority of these polypeptides transport molecules across biological membranes and have been designated the ATP binding cassette (ABC) transporters (Hyde *et al.*, 1990) or traffic ATPases (Ames *et al.*, 1992). A signature pattern for this superfamily of proteins, distinguishing them from other ATPases, consists of the amino acids [LIVMFYC]-[SA][SAPGVYKQ]-G-[DENQMW]-[KRQAPCLW]-[KRNQSTAVM]-[KRACLVM]-[LIVMYPAN]-[PHY]-[LIVMFW]-[SAGCLIVP]-{FYWHP}-KRHP-[LIVMFYWSTA] situated between the paired consensus sequences of the nucleotide binding domain (Bairoch, 1992). This consensus is manifested as the sequence LSG-GQKQRIAIA in virtually all of the P-glycoproteins (Figure 2). In most members of this large group of polypeptides, however, the similarity is more extensive and in many, extends to predicted transmembrane domains. These transport proteins can be categorized into four major subgroups: multicomponent transport systems which include a small hydrophilic polypeptide with a nucleotide binding domain which either 1) shuttle substrates into bacteria, termed the periplasmic permease systems, or 2) export substrates and toxins out of microorganisms, 3) polypeptides approximating one half of a P-glycoprotein with a hydrophobic and hydrophilic domain and 4) transport proteins with a structure similar to the P-glycoprotein comprised of two similar halves each containing transmembrane domains and a nucleotide binding domain (Figure 6). A variety of configurations are capable of transporting similar substrates in different organisms.

The complete sequencing of the yeast *S. cerevisiae* genome has enabled cataloging the entire complement of ABC transport proteins in this organism. Twenty nine members of the superfamily have been identified (Decottignies and Goffeau, 1997). Of note only one ABC transporter, STE6, has high amino acid similarity and topographical organization with the mammalian P-glycoprotein. This is in contrast to all other eukaryotic species which have multigene families of P-glycoprotein homologs when systematically analyzed (Childs and Ling, 1994). On the other hand,

there is a multigene family of nine members which include PDR5 and SNQ2 which are mirror images of the mammalian P-glycoprotein with the nucleotide binding fold followed by the transmembrane domains. In addition, a multigene family comprised of six members, which includes YCF1 and YOR1, has highest similarity to MRP and CFTR has been identified. It is quite interesting that the vast majority of ABC transport proteins in eukaryotic organisms are comprised of proteins which contain transmembrane domains and nucleotide binding folds linked together in comparison with the prokaryotic organisms which have a large number of ABC transporters comprised of multicomponent units, such as found with the bacterial permeases. Analysis of the human expressed tag database indicates that at least 21 distinct ABC transporters remain to be characterized (Allikmets *et al.*, 1996).

Half P-glycoproteins

Polypeptides approximating one half of a P-glycoprotein, consisting of a series of transmembrane loops followed by a hydrophilic region containing a potential nucleotide binding domain, have been identified in both prokaryotes and eukaryotes (Table III and Figure 6). A series of prokaryotic transport proteins which export virulence toxins (HlyB, LtkB, CyaB, ClyI-B, CyaB), polysaccharides (NdvA, ChvA), antibiotics (CvaB, SpaB, LcnC, SyrD, NisT), protease activities (PrtD), heavy metals (Hmt), and the toxic compound aminotriazole (Atr1) have been identified. In addition, prokaryotic polypeptides which are involved in the formation of a polysaccharide layer during heterocyst differentiation (HetA), growth in temperatures greater than 32 °C (MsbA), competence for genetic transformation (ComA) and association with a tRNA protein transferase (MdrA) have been shown to have a similar structure. LmrA exports a broad range of cytotoxic amphiphilic compounds from the lactic acid bacteria *Lactococcus lactis*. When this half P-glycoprotein is transfected into mammalian cells it appears to function as a homodimer capable of conveying multidrug resistance to the same cytotoxic compounds transported by the P-glycoprotein (van Veen *et al.*, 1998). The hydrophobicity plots of each of these polypeptides are very similar to either half of the P-glycoprotein. A comparison of the primary structure indicates an identical amino acid similarity to the P-glycoprotein ranging from 25–35% with an additional 20% conserved residues spanning the entire polypeptide. Similarity is

highest in the hydrophilic domain with 40–45% identical amino acids. Significant nucleic acid similarity, in the range of 45% surrounding the nucleotide binding domain, is even present between the mammalian P-glycoproteins and several of the prokaryotic sequences including *hlyB*, *cyaB*, *msbA*, *lcnC*, *chvA*.

Recently, two mammalian polypeptides, TAP1 and TAP2, have been implicated in the transport of short peptides from the cytosol into the endoplasmic reticulum where they are bound by the major histocompatibility class I molecules for antigen presentation (Townsend and Trowsdale, 1993). Each of these polypeptides appears to represent half of a P-glycoprotein with eight putative transmembrane domains followed by a cytoplasmic domain containing the consensus sequence for a nucleotide binding fold. Comparison of the amino acid sequences of the two peptide transporters indicates approximately 40% amino acid identity with an additional 20% conserved residues. Similarity is highest surrounding the nucleotide binding domain and decreases in the proximal regions. Comparison of the amino acid sequence of the human peptide transporters to human *mdr1* indicates approximately 30% identical and an additional 25% conserved amino acid residues. Both of the peptide transporters are slightly more similar to the carboxy half of the human P-glycoprotein. The genes encoding these peptide transporters are located within 10 kbp of each other in the major histocompatibility class II region. The two polypeptides appear to assemble and function as a heterodimer with both transporters required for the formation of stable class I molecules and antigen presentation (Kelly *et al.*, 1992). Again, the predicted functional structure appears similar to the proposed structure of the P-glycoprotein.

A second group of mammalian half P-glycoproteins are associated with peroxisomes. A partially deleted half P-glycoprotein, ALD, has been identified in some patients with adrenoleukodystrophy (Mosser *et al.*, 1993). The principal abnormality in ALD is the accumulation of very-long-chain fatty acids due to impaired oxidation in peroxisomes. The candidate gene is suspected of translocating a component of the oxidative metabolic pathway into the organelle. This gene has highest similarity, 39% amino acid identity, with another 70 kD peroxisomal membrane protein, HUMPMP, of unknown function (Kamijo *et al.*, 1992). It, too, is presumed to transport substrates into the peroxisome since the nucleotide binding domain is localized to the cytoplasmic side of the organelle. Both peroxisome transporters have approximately 20–

Table 3. Polypeptides with amino acid similarity to the P-glycoprotein. Annotated BLAST search of the SwissProt data base using human *Mdr1* as the query sequence. Results separated into categories described in the text. Original BLAST retrieved 322 entries. Duplicates, similar operons in different species and strains, partial sequences and certain open reading frames (ORF) with unknown functions removed. References available from database designations.

			Smallest Poisson Probability P(N)				Smallest Poisson Probability P(N)
P-glycoproteins				SCRT_DROME	Scarlet		5.0e-10
				WHIT_DROME	White		3.7e-10
MDR1_HUMAN	Mdr1 Human	0.0		ALD_HUMAN	Adrenoleukodystrophy Protein		3.9e-09
MDR1_CRIGR	Pgp1 Hamster	0.0		PMP7_HUMAN	Pmp70 Peroxisome Protein		4.1e-09
MDR3_MOUSE	Mdr3 Mouse	0.0		BROW_DROME	Brown		1.8e-07
MDR2_CRIGR	Mdr2 Hamster	0.0		PAT2_YEAST	Peroxisome Fatty Acid Import		3.0e-05
MDR1_MOUSE	Mdr1 Mouse	0.0					
MDR1_RAT	Mdr1b Rat	0.0		Bacterial Permeases			
MDR3_HUMAN	Mdr3 Human	0.0		REBA_ECOLI	Arabinose Uptake		1.7e-32
MDR2_MOUSE	Mdr2 Mouse	0.0		XYLG_ECOLI	Xylose Uptake		1.6e-27
MDR3_CRIGR	Mdr3 Hamster	0.0		GLNQ_ECOLI	Glutamine Uptake		3.9e-26
MDR2_RAT	Mdr2 Rat	0.0		POTA_ECOLI	Spermidine Uptake		1.1e-23
MDR1_CAEEL	Celpgpa C.elegans	3.0e-288		MGIA_ECOLI	Galactose Uptake		1.4e-23
MDR4_DROME	Mdr49 Drosophila	1.2e-264		PSTB_ECOLI	Phosphate Uptake		1.8e-23
MDR5_DROME	Mdr50 Drosophila	3.1e-256		ARAG_ECOLI	Arabinose Uptake		2.3e-23
MDR3_CAEEL	Celpgpc C.elegans	8.1e-242		OPPD_BACSU	Oligopeptide Uptake		1.3e-22
PMO1_SCHPO	PMO1+ S.Pombe	9.5e-232		PROV_ECOLI	Osmoregulation		1.0e-22
MDR1_LETEN	Lendr Leishmania	5.7e-212		NIKE_ECOLI	Nickel Uptake		9.2e-22
MDR_PLAFF	PMO1 P. Falciparum	2.1e-120		HISP_ECOLI	Histidine Uptake		6.6e-21
STE6_YEAST	STE6	1.3e-116		LACK_AGRD	Lactose Uptake		6.5e-21
MDL_ECOLI	Mdl1	1.0e-87		NOCP_AGR5	Octopine Uptake		4.0e-21
HST6_CANAL	STE6 C.Albicans	5.2e-84		SFUC_SERMA	Iron Uptake		3.1e-21
SUR_CRICR	Sulphonylurea Receptor	1.0e-52		POTG_ECOLI	Putrescine Uptake		1.3e-21
YCF1_YEAST	Cadmium Resistance	1.2e-52		OPUA_BACSU	Glycine Betain Transport		7.0e-21
MRP1_HUMAN	MRP Human	2.1e-51		GLTL_ECOLI	Glutamate/Aspartate Uptake		1.4e-20
YOR1_YEAST	Oligomycin Resistance	1.1e-49		FECE_ECOLI	Ferric Uptake		7.6e-20
CFTR_HUMAN	CFTR	5.0e-39		DPPF_ECOLI	Dipeptide Uptake		8.7e-18
MDR_LEITA	Ltppga Leishmania	3.2e-37		FEPC_ECOLI	Ferric Enterobactin Uptake		8.1e-18
ABC1_MOUSE	ORF Mouse	5.9e-30		MALK_ECOLI	Maltose Uptake		5.6e-18
SNQ2_YEAST	Multiple Drug Resistance	1.4e-13		CYSA_ECOLI	Sulfate Uptake		4.9e-18
CDR1_CANAL	Azole Antifungal Resistance	2.8e-11		BRAG_PSEAE	Branched-chain AA Uptake		5.8e-17
BFR1_SCHPO	Brefeldin A Resistance	3.8e-11		SAPP_HAEIN	Peptide Uptake		7.4e-17
PDR5_YEAST	Pleiotropic Drug Resistance	1.9e-10		MSMK_STRMU	Multiple Sugar Uptake		2.5e-17
Half P-glycoproteins				ARTP_ECOLI	Arginine Uptake		1.3e-17
MSBA_ECOLI	MsbA Growth >32°C	4.6e-82		UGPC_ECOLI	Maltose Uptake		8.2e-16
HETA_ANASP	Heterocyst Differentiation	1.0e-68		P29_MYCHR	M.hyorhins ORF		2.0e-16
TAP1_HUMAN	MHC Peptide Transporter	3.2e-67		LIVF_ECOLI	Branched-chain AA Uptake		7.8e-15
RT1B_ACTPL	RTXI/II Toxin Export	4.9e-66		MODF_ECOLI	Molybdenum Uptake		1.3e-14
HLVB_ECOLI	Hly Toxin Export	2.8e-65		DPPD_ECOLI	Dipeptide Uptake		2.4e-12
TAP2_HUMAN	MHC Peptide Transporter	1.1e-62		NOSF_PSEST	Copper Uptake		2.3e-12
TAGB_DICDI	Serine Protease	1.8e-59		PHNK_ECOLI	Alkylphosphonates Uptake		1.1e-11
CHVA_AGRU	Beta-(1-->2)Glucan Export	1.3e-58		NIKD_ECOLI	Nickel Uptake		7.1e-08
NDVA_RHIME	Beta-(1-->2)Glucan Export	1.6e-58		BTUD_ECOLI	Vitamin B12 Uptake		1.9e-07
ATM1_YEAST	Mitochondrial Transporter	5.6e-58		TAGH_BACSU	Teichoic Acid Uptake		3.1e-05
HMT1_SCHPO	Heavy Metal Resistance	2.6e-56		Multicomponent Exporters			
CYAB_BORPE	Cya Toxin Export B.pertusis	3.2e-55		NODI_RHILV	β-1,4 Glucosamine Export		8.9e-19
COMA_STRPN	Competence Induction	2.7e-46		DRRA_STRPE	Daunorubicin Resistance		2.5e-12
PEDD_PEDAC	Bacteriocin Export	2.3e-45		TLRC_STRFR	Tylosin Resistance		5.8e-10
LCNC_LACLA	Lactococcin A Export	1.4e-45		CCMA_ECOLI	Heme Export		6.8e-11
CVAB_ECOLI	ColicinV Export	2.7e-38		MSRA_STAEP	Erythromycin Resistance		5.8e-09
CYDC_ECOLI	Cytochrome D Export	1.4e-39		BEXA_HAEIN	Polysacchride to Capsule		7.4e-08
SPAB_BACSU	Lantibiotic Subtilin Export	3.0e-33		KST5_ECOLI	Polysialic Acid to Capsule		9.0e-05
APRD_PSEAE	Alkaline Phosphatase Export	2.3e-33					
PRTD_ERWCH	Protease Export	2.2e-30		Other			
LCN3_LACLA	Lantibiotic Lactacin Export	2.7e-26		FTSE_ECOLI	Cell Division Protein		2.8e-17
NIST_LACLA	Nicin Export	3.5e-24		EF3H_YEAST	Elongation Factor3		7.7e-13
WHIT_HUMAN	White Human	7.7e-14		UVRA_ECOLI	DNA Excision Repair		6.1e-09
SYRD_PSESY	Syringomycin Synthesis	4.3e-14					
ADP1_YEAST	ORF with EGF Motif	1.2e-13					

25% amino acid identity with the peptide transporters as well as human MDR1, most of which surrounds the nucleotide binding domain.

Disruption of the prestalk specific TagB gene in *Dictyostelium* results in a block in development and differentiation (Shaalsky *et al.*, 1995). This ABC transporter family member is arranged like a half P-glycoprotein, but has an additional 1000 amino acids at the N-terminus which include sequence similarity to a family of serine proteases. In addition, partial sequence of TagC suggests a polypeptide with high similarity to TagB and a similar phenotype when disrupted, suggesting possible functioning as a heterodimer.

The *Drosophila* white, brown and scarlet gene products are structurally related to the P-glycoprotein. Each consists of a nucleotide binding fold followed by a series of transmembrane domains (Dreesen *et al.*, 1988; O'Hare *et al.*, 1984). The products of these genes have been proposed to transport pigment precursors in the eye, possibly functioning as heterodimers. Human homologs with high similarity to the *Drosophila* protein have been identified (Chen *et al.*, 1996; Croop *et al.*, 1997; Savary *et al.*, 1996). They have highest homolog similarity with the recently identified *Drosophila* ATET protein which is found in the bacterial permeases originally had significant latitude as building blocks in forming transport proteins (Figure 6).

Bacterial permeases

The bacterial periplasmic permeases provide an energy dependent transport system for the uptake of nutrients in Gram negative bacteria (Ames *et al.*, 1990; Higgins *et al.*, 1990). These multicomponent systems provide a mechanism for specific sugars, metals, amino acids, peptides, or co-factors which have traversed the outer membrane into the periplasmic space to cross the inner cell membrane (Table III and Figure 6). The permeases are characteristically comprised of four individual polypeptides encoded in one or two operons. Each individual system includes a periplasmic substrate-binding protein which is specific for the substrate. Although there is little primary amino acid similarity amongst the substrate-binding proteins, a striking similarity in the tertiary structure has been deduced by x-ray crystallography (Adams and Oxender, 1989). When bound to the substrate, the binding proteins undergo a conformational change which allows interaction with the membrane bound components of the system. The membrane bound components of the

transport system are comprised of two hydrophobic polypeptides and a third member with a nucleotide binding domain. This latter polypeptide displays a high level of similarity to the analogous member of each permease as well as the entire cytoplasmic domain of the P-glycoprotein. The nucleic acid sequence encoding the region around the nucleotide binding domain even displays nucleic acid similarity to the mammalian P-glycoproteins in some of the permeases such as *potA* of the spermidine and putrescine transport system, and *braG* of the branched chain amino acid transport system (data not shown). The two hydrophobic polypeptides are each comprised of five or six putative transmembrane domains with approximately 25–30% identical and an additional 25% conserved amino acid homology with each other. In a high affinity *Mycoplasma* transport system, the hydrophobic proteins appear to be fused together as a single polypeptide which is consistent with the proposed function of the hydrophobic domains as pseudodimers (Higgins *et al.*, 1990). Some of the hydrophobic domains in different permeases display homology to each other and in some cases (*upgE*, *araJ*, *oppB*) even very faint homology to the P-glycoprotein (data not shown). Although the domain with homology to the cytoplasmic region of P-glycoprotein does not include a large number of hydrophobic residues, it appears to be membrane bound on the cytoplasmic side of the inner membrane (Ames, 1986). In several of the permease systems this domain with the nucleotide binding fold is duplicated (*mglA*, *rbsA*, *araG*) and in the *opp* system there are two similar polypeptides with nucleotide binding domains, both of which are required for transport. ATP hydrolysis is tightly coupled to the transport of the substrates and presumably results in conformational changes which allow translocation of the substrate across the membrane. The entire permease complex, comprised of polypeptides containing twelve membrane spanning regions associated with two non-hydrophobic polypeptides which include nucleotide binding domains, suggests a functional unit quite similar to the predicted structure of the P-glycoprotein.

Multicomponent export systems

Hydrophilic polypeptides with homology to the cytoplasmic region of the P-glycoprotein act as members of multi-component transport systems involved in the export of a variety of substrates from microorganisms (Table III and Figure 6). *Streptomyces* which pro-

duce doxorubicin and daunorubicin express two genes located in tandem which convey resistance to these antibiotics in heterologous hosts (Guilfoile and Hutchinson, 1991). One of the polypeptides is hydrophilic and contains a nucleotide binding domain, DrrA, and the other hydrophobic with six potential transmembrane domains, TrrB. Another resistance operon in *Streptomyces* encodes OleC, a polypeptide comprised of a nucleotide binding domain, and a second hydrophobic polypeptide with six putative transmembrane domains. Together these polypeptides encode resistance to oleandomycin. Similarly, polypeptides containing two nucleotide binding domains linked in tandem have been identified in *Streptomyces* which produce macrolide antibiotics (Schoner *et al.*, 1992). These polypeptides act as specific resistance determinants in *Streptomyces* producing spiramycin, SrmB, carbomycin, CarA, and tylosin, TlrC, and appear to transport the antibiotics out of the organisms. Other potential members of the transport systems have not yet been identified.

Bacterial operons responsible for the export of antibiotics have also been identified. Host resistance to the peptide antibiotic microcin B17 is encoded in the *mbcF* gene in *Enterobacteriaceae*. The MbcF polypeptide includes a single nucleotide binding domain and interacts with a hydrophobic polypeptide to export microcin B17 from the cytoplasmic to the periplasmic space (Garrido *et al.*, 1988). The biosynthesis operon of the peptide, lantibiotic epidermin from *Staphylococci*, similarly encodes a polypeptide with a nucleotide binding domain presumably responsible for export (Schnell *et al.*, 1992). Polypeptides consisting of duplicated nucleotide binding domains linked together encode an inducible mechanism of erythromycin resistance, MsrA, and virginiamycin resistance, VgA, in *Staphylococci*. MsrA has been shown to function by decreasing accumulation of the antibiotic (Ross *et al.*, 1990). The hydrophilic and hydrophobic domains of these multicomponent units presumably interact to transport the toxic antibiotics out of the organisms. Although, the structural organization of the efflux complexes have not yet been clarified, the similarity of the components to those found in bacterial permeases suggest that the two may be similar. Understanding the mechanism of substrate influx versus toxin efflux will certainly provide insight into the transport mechanisms of the P-glycoprotein.

Multicomponent transport systems responsible for the transport of substrates which form the bacterial capsule and are responsible for microbial viru-

lence have similarly been identified. The *bexA* gene encodes a hydrophilic polypeptide containing a nucleotide binding domain and *bexB* a polypeptide with six potential transmembrane domains in *H. influenzae* type b (Kroll *et al.*, 1988). Both are essential for polysaccharide export to the capsule and disruption of this complex results in loss of capsule formation. The polysaccharide capsule of *N. meningitidis* requires the function of a similar operon, *ctr*. Similarly, *kpsT* and *kpsM* encode analogous proteins which are required for polysialic acid incorporation in the capsule of *E. coli* K1, the causative agent of neonatal septicemia and meningitis. These transport systems appear to represent a generalized mechanism to translocate substrates across the cytoplasmic membrane for insertion in the capsule. Knowledge of these systems in pathogenic bacteria provide a new target for antibiotics by inhibition of substrate transport, analogous to the action of reversal agents in multidrug resistant cells.

Several systems which may not transport substrates also include hydrophilic domains with homology to P-glycoprotein. The *ftsE* gene is required by *E. coli* for cellular division. Within this operon is an open reading frame for a hydrophilic polypeptide which contains the consensus sequence for a nucleotide binding domain (Gill *et al.*, 1986). The ABC excinuclease of *E. coli*, an ATP dependent DNA repair enzyme, is made up of three subunits, including UvrA with a structure comprised of a duplication with nucleotide binding domains in each half (Husain *et al.*, 1986). Similarly, elongation factor-3 of *S. cerevisiae*, which appears to transduce energy for stimulation of aminoacyl tRNA binding to the ribosome, has a duplicated structure with a nucleotide binding domain in each half (Qin *et al.*, 1990). These diverse examples suggest that the functional unit used to bind and hydrolyze ATP in the transport systems can be utilized for ATP hydrolysis in a variety of cellular functions. This observation has further suggested that the hydrophobic domains of the transport complexes are responsible for substrate specificity and translocation.

Evolutionary origins of the P-glycoprotein

The P-glycoprotein is a member of a large superfamily of transport proteins which move a wide range of substrates across plasma membranes. The general structure of 12 membrane spanning domains and two hydrophilic moieties which include nucleotide bind-

ing domains appears similar across large evolutionary distances. However, the composition of the functional units comprising the structure varies and a range of combinations can direct transport in and out of the cell as well as across different organelles. Cytotoxic compounds are exported not only by P-glycoproteins, but also by multicomponent transport systems and half P-glycoproteins in a variety of organisms.

Although the functional domains of the P-glycoprotein are observed in both prokaryotes and eukaryotes, the precise evolutionary origins have been a matter of speculation. The observation that the two halves of the P-glycoprotein displayed a high level of amino acid homology originally suggested that the polypeptide represented gene duplication of a primordial gene. Identification of the intron/exon boundaries of the human *mdr1* (Chen *et al.*, 1990), *mdr3* (Lincke *et al.*, 1991) and mouse *mdr1* (Raymond and Gros, 1989) P-glycoproteins have provided more information on the possible evolutionary mechanisms involved. Each of the mammalian homologs contains 28 exons with introns placed at homologous sites along the gene. The most notable feature is that the majority of the boundaries in the proximal and distal halves of the gene do not occur at the corresponding locations in the two halves. Only two intron boundaries are located at identical nucleotides in the two halves and occur within the nucleotide binding domain. Another intron boundary in a hydrophobic region occurs at the same amino acid. The lack of similarity between the halves suggested that the P-glycoprotein may have originated from two genes with common ancestors evolving separately prior to fusion (Chen *et al.*, 1990). However, since both moieties have a similar number of exons, over 50% nucleic acid identity in the coding regions and the exons appear to encode functional domains, the theory for gene duplication has similarly been advanced (Raymond and Gros, 1989). Further insight has been provided with determination of the intron/exon boundaries of the *A. thaliana* homolog (Dudler and Hertig, 1992). Nine introns have been identified, eight in the proximal half and one in the distal half. None of the introns in the proximal half match the position of the intron in the distal half. Four of the introns correspond to similar positions in the mammalian gene including two in the proximal half aligning with the introns which are identical in the nucleotide binding domain. In addition, however, three introns in one half of the *A. thaliana* homolog correspond to introns in the opposite half of the mammalian gene. Preliminary observations in the

C. elegans (Lincke *et al.*, 1992) and *Drosophila* homologs (J. Croop, unpublished data) suggest a similar conservation of introns in one half of the homolog with introns in the contralateral half of the mammalian gene. These latter observations suggest that a duplicated primordial gene could have differentially lost introns to account for the variability between the proximal and distal halves of the mammalian P-glycoprotein. This would also explain the similarity found in intron location across different halves in the evolutionarily divergent homologs (Dudler and Hertig, 1992). Such a scenario, however, might be expected to lead to exons with larger than average size in the mammalian gene, which is not the case (Chen *et al.*, 1990).

Examination of the similarity among the currently described P-glycoproteins reveals several distinct groupings which presumably have evolutionary significance. The mammalian P-glycoproteins can be identified as a single group with over 65% identity amongst all of the homologs as indicated in Figures 6 and 7. Within each species, the mammalian P-glycoproteins display very high levels of similarity in the range of 72–85%. A similar level of homology is present between the strain specific *lemdr1* and *ldmdr1*, but is not found among the homologs in other organisms. The two *E. histolytica* approach this level of homology with 68% identity, but the homologs in *Drosophila* and *C. elegans* are as divergent from each other as they are from the other P-glycoproteins. Interestingly, most of the homologs, except for the ones noted above, display a slightly higher homology to a mammalian P-glycoprotein than any other homolog (data not shown). The recently described clustering of homologies with *Mrp* and *CFTR* can similarly be seen in Figure 7 (Cole *et al.*, 1992). However, note that the amino acid identities among these three homologs are only 20–30% compared to approximately 15% identity with each of the other P-glycoprotein homologs. The majority of the remaining homologs display over 30% amino acid identity to one another. The exceptions are *pfmdr1*, *Ste6*, and *Mdl* which appear to represent a separate grouping with only 20–30% amino acid identity with the other homologs. These four groups are similarly distinguished in the pairings found in Figure 7.

Direct comparisons of the P-glycoprotein homologs to human MDR1 support these groupings. Figure 4A clearly indicates the high level of homology to the mammalian P-glycoproteins, a lesser but similar level of homology to each of the remaining

homologs with the exception of STE6 and MDL and MRP, LTPGP1 and CFTR which segregate into two discrete groups. The PFMDR1 homolog is intermediate in homology to STE6 and MDL and the majority of homologs. An similar result is obtained comparing *Drosophila* MDR49 to the other homologs (Figure 4B). In addition, a BLAST search of the genetic databases (Altshul *et al.*, 1990) to identify proteins with homology to human MDR1 indicates a similar clustering of the homologs. Table III indicates that the probability approaches zero for a chance occurrence to have identified human MDR1 and the other mammalian homologs. The other homologs are clustered as described above with PFMDR1, STE6, and MDL AND LTPGP1, MRP, and CFTR, and the remainder of the homologs defining three distinct levels in the probability of being similar to human MDR1. These distinctions are highlighted by the fact that many of the half P-glycoproteins are more likely than PFMDR1, STE6, MDL, LTPGP1, MRP, or CFTR to have matched human *mdr1*. Furthermore, BLAST searches of the genetic databases with nucleic acid sequences from the nucleotide binding domains of the mammalian P-glycoproteins identify all of the homologs except STE6, MDL, LTPGP1 and CFTR (data not shown). However, a number of half P-glycoproteins and multicomponent transport genes are identified. Finally, BLAST searches with the nucleic acid sequences from the carboxy half of the mammalian P-glycoproteins consistently match many more of the homologs as well as other polypeptides with evolutionary homology than searches using the amino end as the query sequence.

Characterization of the differences in homology between the proximal and distal halves of the P-glycoprotein homologs is also consistent with the assignment of these groupings. The mammalian P-glycoproteins display the highest levels of homology between the two halves with approximately 40% identical amino acid residues. The majority of the remaining homologs are 30–40% identical, while LTPGP1, MRP, and CFTR are 21–23% identical. The two halves of PFMDR1, STE6, and MDL range from 24 to 29% identity. Differential levels of similarity between the two halves suggest that the structural or functional requirements have evolved over time. Unfortunately, the normal physiological roles for most P-glycoproteins are unknown so it is difficult to determine the significance of these differences. Since, in general, the carboxy and amino halves are more similar to the respective halves among P-glycoprotein homologs (data not

shown), either similarity remains to the original ancestral moieties which fused to form the P-glycoprotein or divergence between the two halves of a duplicated gene occurred prior to further radiation. These comparisons have started to delineate distinct groupings which appear to correlate with differences in the specificity of function. Identification of the complete set of P-glycoproteins and ABC transporters should provide valuable clues to structure-function relationships and evolutionary origins.

Acknowledgements

This work was supported in part by a grant from Sandoz Pharmaceuticals.

References

- Abraham EH, Prat AG, Gerweck L, Seneveratne T, Arceci R, Kramer R, Guidotti G and Cantiello HF (1993) The multidrug resistance (*mdr1*) gene product functions as an ATP channel. *Proc Natl Acad Sci* 90: 312–316.
- Adams MD and Oxender DL (1989) Bacterial periplasmic binding protein tertiary structures. *J Biol Chem* 264: 15739–15742.
- Alarco AM, Balan I, Talibi D, Mainville N and Raymond M (1997) AP1-mediated multidrug resistance in *Saccharomyces cerevisiae* requires FLR1 encoding a transporter of the major facilitator superfamily. *J Biol Chem* 272: 19304–19313.
- Allikmets R, Gerrard B, Court D and Dean M (1993) Cloning and organization of the *abc* and *mdl* genes of *Escherichia coli*: relationship to eukaryotic multidrug resistance. *Gene* 136: 231–236.
- Allikmets R, Gerrard B, Hutchinson A and Dean M (1996) Characterization of the human ABC superfamily: isolation and mapping of 21 new genes using the expressed sequence tags database. *Hum Mol Gen* 5: 1649–1655.
- Altshul SF, Gish W, Miller W, Meyers E and Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.
- Ambudkar SV, Lelong IH, Zhang J, Cardarelli CO, Gottesman Mm and Pastan I (1992) Partial purification and reconstitution of the human multidrugresistance pump: characterization of the drug-stimulatable ATP hydrolysis. *Proc Natl Acad Sci* 89: 8472–8476.
- Ames G (1986) Bacterial periplasmic transport systems: structure, mechanism and evolution. *Annu Rev Biochem* 55: 397–425.
- Ames GF, Mimura CS, Holbrook SR and Shyamala V (1992) Traffic ATPases: a superfamily of transport proteins operating from *Escherichia coli* to humans. *Adv Enzym and Rel Areas of Mol Biol* 65: 1–47.
- Ames GF, Mimura CS and Shyamala V (1990) Bacterial periplasmic permeases belong to a family of transport proteins operating from *Escherichia coli* to human: Traffic ATPases. *Fems Microbiol Rev* 6: 429–446.
- Arceci RJ, Baas F, Raponi R, Horwitz SB, Housman D and Croop J (1990) Multidrug resistance gene expression is controlled by steroid hormones in the secretory epithelium of the uterus. *Mol Reprod Dev* 25: 101–109.
- Arceci RJ, Croop JM, Horwitz SB and Housman D (1988) The gene encoding multidrug resistance is induced and expressed at high

- levels during pregnancy in the secretory epithelium of the uterus. *Proc Natl Acad Sci* 85: 4350–4354.
- Baas F and Borst P (1988) The tissue dependent expression of hamster P-glycoprotein genes. *FEBS Letters* 229: 329–332.
- Baichwal V, Liu D and Ames GF (1993) The ATP-binding component of a prokaryotic traffic ATPase is exposed to the periplasmic (external) surface. *Proc Natl Acad Sci* 90: 620–624.
- Bairoch A (1992) PROSITE: a dictionary of sites and patterns in proteins. *Nucleic Acids Res* 20, Suppl: 2013–2018.
- Balzi E and Goffeau A (1995) Yeast multidrug resistance: the PDR network. *J Bioenergetics Biomembranes* 27: 71–76.
- Beck WT (1990) Multidrug resistance and its circumvention. *Eur J Cancer* 26: 513–515.
- Beja O and Bibi E (1995) Multidrug resistance protein (Mdr)-alkaline phosphatase hybrids in *Escherichia coli* suggest a major revision in the topology of the C-terminal half of Mdr. *Journal of Biological Chemistry* 270: 12351–12354.
- Berkower C and Michaelis S (1991) Mutational analysis of the yeast a-factor transporter STE6, a member of the ATP binding cassette (ABC) protein superfamily. *Embo J* 10: 3777–3785.
- Berkower C, Taglicht D and Michaelis S (1996) Functional and physical interactions between partial molecules of STE6, a yeast ATP-binding cassette protein. *J Biol Chem* 271: 22983–22989.
- Biedler JL and Riehm H (1970) Cellular resistance to actinomycin D in Chinese hamster cells in vitro: cross-resistance, radioautographic and cytogenetic studies. *Cancer Res* 30: 1174–1184.
- Bosch I and Croop JM (1996) P-glycoprotein multidrug resistance and cancer. *Biochim Biophys Acta* 1288: F37–F54.
- Bosch I, Dunussi-Joannopoulos K, Wu RL, Furlong S and Croop JM (1997) Phosphatidylcholine and Phosphatidylethanolamine are substrates of human MDR1 P-glycoprotein. *Biochem (in press)*.
- Bosch I, Jackson GR, Croop JM and Cantiello HF (1996) Expression of *Drosophila melaogaster* P-glycoproteins is associated with ATP-channel activity. *Am J Physiol* 40: C1527–C1538.
- Bosch I, Wang Z, Tao L and Shoemaker C (1994) Two *Schistosoma Manonsi* cDNAs encoding ATP binding cassette (ABC) family proteins. *Mol Biochem Parasit* 65: 351–356.
- Boscoboinik D, Debanne MT, Stafford AR, Jung CY, Gupta R and Epand RM (1990) Dimerization of the P-glycoprotein in membranes. *Biochim Biophys Acta* 1027: 225–258.
- Bradley G, Georges E and Ling V (1990) Sex-dependent and independent expression of the P-glycoprotein isoforms in chinese hamster. *J Cell Physiol* 145: 398–408.
- Broeks A, Janssen HW, Calafat J and Plasterk RH (1995) A P-glycoprotein protects *Caenorhabditis elegans* against natural toxins. *Embo J* 14: 1858–1866.
- Buschman E, Arceci R, Croop J, Mingxin C, Arias I, Housman D and Gros P (1992) Isoform specific antibodies identify the bilie canalicular membrane as the primary site of expression of the P-glycoprotein encoded by mouse *mdr2*. *J Biol Chem* 267: 18093–18099.
- Callahan HL and Beverley S (1991) Heavy metal resistance: a new role for P-glycoproteins in *Leishmania*. *J Biol Chem* 266: 18427–18430.
- Castillo G, Shen HJ and Horwitz SB (1995) A homologue of the mammalian multidrug resistance gene (*mdr*) is functionally expressed in the intestine of *Xenopus laevis*. *Biochim Biophys Acta* 1262: 113–123.
- Chan KM, Davies PL, Childs S, Veinot L and Ling V (1992) P-glycoprotein genes in the winter flounder, *Pleuronectes americanus*: isolation of two type of genomic clones carrying 3' terminal exons. *Biochim Biophys Acta* 1171: 65–72.
- Chaudhary PM, Mechetner EB and Roninson IB (1992) Expression and activity of the multidrug resistance P-glycoprotein in human peripheral blood lymphocytes. *Blood* 80: 2735–2739.
- Chaudhary PM and Roninson IB (1991) Expression and activity of P-glycoprotein, a multidrug efflux pump, in human hematopoietic stem cells. *Cell* 66: 85–94.
- Chaudhary PM and Roninson IB (1993) Induction of multidrug resistance in human cells by transient exposure to different chemotherapeutic drugs. *J Natl Cancer Inst* 85: 632–639.
- Chen CJ, Chin JE, Ueda K, Clark DP, Pastan I, Gottesman Mm and Roninson IB (1986) Internal duplication and homology with bacterial transport proteins in the *mdr1* (P-glycoprotein) gene from multidrug-resistant human cells. *Cell* 47: 381–389.
- Chen CJ, Clark D, Ueda K, Pastan I, Gottesman Mm and Roninson IB (1990) Genomic organization of the human multidrug resistance (MDR1) gene and origin of P-glycoproteins. *J Biol Chem* 265: 506–514.
- Chen H, Rossier C, Lalioti MD, Lynn A, Chakravarti A, Perrin G and Antonarakis SE (1996) Cloning of the cDNA for a human homologue of the *Drosophila* white gene and mapping to chromosome 21q22.3. *American Journal of Human Genetics* 59: 66–75.
- Childs S and Ling V (1994). The MDR superfamily of genes and its biological implications. *Important Advances in Oncology*. V. DeVita, S. Hellman and S. Rosenberg. Philadelphia, Lippincott: 21–36.
- Childs S and Ling V (1996) Duplication and evolution of the P-glycoprotein genes in pig. *Biochim Biophys Acta* 1307: 205–212.
- Childs S, Yeh RL, Georges E and Ling V (1995) Identification of a sister gene to P-glycoprotein. *Cancer Res* 55: 2029–2034.
- Chin JE, Soffir R, Noonan KE, Choi K and Roninson IB (1989) Structure and expression of the human MDR (P-glycoprotein) gene family. *Mol Cell Biol* 9: 3808–3820.
- Choi K, Chen CJ, Krieger M and Roninson IB (1988) An altered pattern of cross-resistance in multidrug-resistant human cells results from spontaneous mutations in the *mdr1* (P-glycoprotein) gene. *Cell* 53: 519–529.
- Chow LM, Wong AK, Ullman B and Wirth DF (1993) Cloning and functional analysis of an extrachromosomally amplified multidrug resistance-like gene in *Leishmania enriettii*. *Mol Biochem Parasitol* 60: 195–208.
- Cianfriglia M, Poloni F, Signoretti C, Romagnoli G, Tombesi M and Felici F (1996) Topology of MDR1-P-glycoprotein as indicated by epitope mapping of monoclonal antibodies to human MDR cells. *Cytotechnology* 19: 247–251.
- Cole SPC, Bhardwaj G, Gerlach JH, Mackie JE, Grant CE, Almquist KC, Stewart AJ, Kurz EU, Duncan AMV and Deeley RG (1992) Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line. *Science* 258: 1650–1654.
- Cordon-Cardo C, OBrien JP, Boccia J, Casals D, Bertino JR and Melamed MR (1990) Expression of the multidrug resistance gene product (P-glycoprotein) in human normal and tumor tissues. *J Histochem Cytochem* 38: 1277–1287.
- Cordon-Cardo C, OBrien JP, Casals D, Rittman-Grauer L, Biedler JL, Melamed MR and Bertino JR (1989) Multidrug-resistance gene (P-glycoprotein) is expressed by endothelial cells at blood-brain barrier sites. *Proc Natl Acad Sci* 86: 695–698.
- Cowman AF, Karcz S, Galatis D and Culvenor JG (1991) A P-glycoprotein homologue of *Plasmodium falciparum* is localized to the digestive vacuole. *J Cell Biol* 113: 1033–1042.
- Croop J, Guild B, Gros P and Housman D (1987) Genetics of multidrug resistance: relationship of a cloned gene to the complete multidrug resistance phenotype. *Cancer Res* 47: 5982–5988.

- Croop JM, Raymond M, Haber D, Devault A, Arceci RJ, Gros P and Housman DE (1989) The three mouse multidrug resistance (*mdr*) genes are expressed in a tissue-specific manner in normal mouse tissues. *Mol Cell Biol* 9: 1346–1350.
- Croop JM, Tiller GE, Fletcher JA, Lux ML, Raab E, Goldenson D, Son D, Arciniegas S and Wu RL (1997) Isolation and Characterization of a mammalian homolog of the *Drosophila* white gene. *Gene* 185: 77–85.
- Cui Z, Hirata D, Tsuchiya E, Osada H and Miyakawa T (1996) The multidrug resistance-associated protein (MRP) subfamily (Yrs1/Yor1) of *Saccharomyces cerevisiae* is important for the tolerance to a broad range of organic anions. *J Biol Chem* 271: 14712–14716.
- Dano K (1973) Active outward transport of daunomycin in resistant Ehrlich ascites tumor cells. *Biochim Biophys Acta* 323: 446–483.
- De Bruijn MH, Van der Blik AM, Biedler JL and Borst P (1986) Differential amplification and disproportionate expression of five genes in three multidrug-resistant Chinese hamster lung cell lines. *Mol Cell Biol* 6: 4717–22.
- Decottignies A and Goffeau A (1997) Complete inventory of the yeast ABC proteins. *Nature Genetics* 15: 137–145.
- Descoteaux S, Ayala P, Orozco E and Samuelson J (1992) Primary sequence of two P-glycoprotein genes of *Entamoeba histolytica*. *Mol Biochem Parasit* 54: 201–211.
- Devault A and Gros P (1990) Two members of the mouse *mdr* gene family confer multidrug resistance with overlapping but distinct drug specificities. *Mol Cell Biol* 10: 1652–1663.
- Devereux J, Haerberli P and Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. *Nuc Acids Res* 12: 387–395.
- Drach D, Zhao S, Drach J, Mahadevia R, Gattringer C, Huber H and Andreeff M (1992) Subpopulations of normal peripheral blood and bone marrow cells express a functional multidrug resistant phenotype. *Blood* 80: 2729–2734.
- Dreesen TD, Johnson DH and Henikoff S (1988) The brown protein of *Drosophila melanogaster* is similar to the white protein and to components of active transport complexes. *Mol Cell Biol* 8: 5206–5215.
- Dudler R and Hertig C (1992) Structure of an *mdr*-like gene from *Arabidopsis thaliana*. *J Biol Chem* 267: 5882–5888.
- Dunia I, Smit JJ, van der Valk MA, Bloemendal H, Borst P and Benedetti EL (1996) Human multidrug resistance 3-P-glycoprotein expression in transgenic mice induces lens membrane alterations leading to cataract. *J Cell Biol* 132: 701–716.
- Duraisingh MT, Drakeley CJ, Muller O, Bailey R, Snounou G, Targett GA, Greenwood BM and Warhurst DC (1997) Evidence for selection for the tyrosine-86 allele of the *pfmdr* 1 gene of *Plasmodium falciparum* by chloroquine and amodiaquine. *Parasitology* 11: 205–211.
- Edgar R and Bibi E (1997) MdfA, an *Escherichia coli* multidrug resistance protein with an extraordinarily broad spectrum of drug recognition. *J Bact* 179: 2274–2280.
- Ellenberger TE and Beverley SM (1989) Multiple drug resistance and conservative amplification of the H region in *Leishmania major*. *J Biol Chem* 264: 15094–15103.
- Endicott J, Juranka P, Farida S, Gerlach J, Deuchars K and Ling V (1987) Simultaneous expression of two P-glycoprotein genes in drug-sensitive Chinese hamster ovary cells. *Mol Cell Biol* 7: 4075–4081.
- Endicott JA and Ling V (1989) The biochemistry of P-glycoprotein-mediated multidrug resistance. *Annu Rev Biochem* 58: 137–171.
- Endicott JA, Sarangi F and Ling V (1991) Complete cDNA sequences encoding the Chinese hamster P-glycoprotein gene family. *DNA Seq* 2: 89–101.
- Felsenstein J (1989) PHYLIP - Phylogeny Inference Package. *Cladistics* 5: 164–166.
- Fojo AT, Ueda K, Slamon DJ, Poplack DG, Gottesman Mm and Pastan I (1987) Expression of a multidrug-resistance gene in human tumors and tissues. *Proc Natl Acad Sci* 84: 265–269.
- Foot SJ, Kyle DE, Martin RK, Oduola AM, Forsyth K, Kemp DJ and Cowman AF (1990) Several alleles of the multidrug-resistance gene are closely linked to chloroquine resistance in *Plasmodium falciparum*. *Nature* 345: 255–258.
- Foot SJ, Thompson JK, Cowman AF and Kemp DJ (1989) Amplification of the multidrug resistance gene in some chloroquine-resistant isolates of *P. falciparum*. *Cell* 57: 921–930.
- Ford JM and Hait WN (1990) Pharmacology of drugs that alter multidrug resistance in cancer. *Pharmacol Rev* 42: 155–199.
- Garrido M, Herrero M, Kolter R and Moreno F (1988) The export of the DNA replication inhibitor Microcin B17 provides immunity for the host cell. *Embo J* 7: 1853–1862.
- Georges E, Bradley G, Garipey J and Ling V (1990) Detection of P-glycoprotein isoforms by gene-specific monoclonal antibodies. *Proc Natl Acad Sci* 87: 152–156.
- Gerrard B, Steward C and Dean M (1993) Analysis of Mdr50: A *Drosophila* P-glycoprotein/multidrug resistance gene homolog. *Genomics* 17: 83–88.
- Ghosh SK, Lohia A, Kumar A and Samuelson J (1996) Overexpression of P-glycoprotein gene 1 by transfected *Entamoeba histolytica* confers emetine-resistance. *Mol Biochem Parasitol* 82: 257–260.
- Gill DR, Hatful GF and Salmond GP (1986) A new cell division operon in *Escherichia coli*. *Mol Gen Genet* 86: 134–145.
- Gottesman M and Pastan I (1989) The multidrug resistance transporter, a double-edged sword. *J Biol Chem* 263: 12163–12166.
- Greenberger LM, Lothstein L, Williams SS and Horwitz SB (1988a) Distinct P-glycoprotein precursors are overproduced in independently isolated drug-resistant cell lines. *Proc Natl Acad Sci* 85: 3762–3766.
- Greenberger LM, Williams SS, Georges E, Ling V and Horwitz SB (1988b) Electrophoretic analysis of P-glycoproteins produced by mouse J774.2 and Chinese hamster ovary multidrug-resistant cells. *J Natl Cancer Inst* 80: 506–510.
- Gros P, Ben Neriah Y, Croop JM and Housman DE (1986a) Isolation and expression of a complementary DNA that confers multidrug resistance. *Nature* 323: 728–731.
- Gros P, Croop J and Housman D (1986b) Mammalian multidrug resistance gene: complete cDNA sequence indicates strong homology to bacterial transport proteins. *Cell* 47: 371–380.
- Gros P, Croop J, Roninson I, Varshavsky A and Housman DE (1986c) Isolation and characterization of DNA sequences amplified in multidrug-resistant hamster cells. *Proc Natl Acad Sci* 83: 337–341.
- Gros P, Dhir R, Croop J and Talbot F (1991) A single amino acid substitution strongly modulates the activity and substrate specificity of the mouse *mdr1* and *mdr3* drug efflux pumps. *Proc Natl Acad Sci* 88: 7289–7293.
- Gros P, Raymond M, Bell J and Housman D (1988) Cloning and characterization of a second member of the mouse *mdr* gene family. *Mol Cell Biol* 8: 2770–2778.
- Guilfoile PG and Hutchinson CR (1991) A bacterial analog of the *mdr* gene of mammalian tumor cells is present in *Streptomyces peucetius*, the producer of daunorubicin and doxorubicin. *Proc Natl Acad Sci* 88: 8553–8557.

- Hardy SP, Goodfellow HR, Valverde MA, Gill DR, Sepulveda FV and Higgins CF (1994) Protein kinase C-mediated phosphorylation of the human multidrug resistance P-glycoprotein regulates cell volume-activated chloride channels. *EMBO J* 14: 68–75.
- Henderson DM, Sifri CD, Rodgers M, Wirth DF, Hendrickson N and Ullman B (1992) Multidrug resistance in *Leishmania donovani* is conferred by amplification of a gene homologous to the mammalian *mdr1* gene. *Mol Cell Biol* 12: 2855–2865.
- Higgins CF, Hyde SC, Mimmack Mm, Gileadi U, Gill DR and Gallagher MP (1990) Binding protein-dependent transport systems. *J Bioenerg Biomembr* 22: 571–592.
- Horio M, Gottesman Mm and Pastan I (1988) ATP-dependent transport of vinblastine in vesicles from human multidrug-resistant cells. *Proc Natl Acad Sci* 85: 3580–3584.
- Hsu S, Lothstein L and Horwitz S (1989) Differential overexpression of the three *mdr* gene family members in multidrug resistant J774.2 mouse cells. *J Biol Chem* 264: 12053–12062.
- Husain I, Van Houten B, Thomas D and Sancar A (1986) Sequences of *Escherichia coli uvrA* gene and protein reveal two potential ATP binding sites. *J Biol Chem* 261: 4895–4901.
- Hyde S, Emsley P, Hartshorn M, Mimmack M, Gileadi U, Pearce S, Gallagher M, Gill D, Hubbard R and Higgins C (1990) Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. *Science* 346: 362–365.
- Jan LY and Jan YN (1992) Tracing the roots of ion channels. *Cell* 69: 715–718.
- Juliano RL and Ling V (1976) A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochim Biophys Acta* 455: 152–162.
- Juranka PF, Zastawny RL and Ling V (1989) P-glycoprotein: multidrug-resistance and a superfamily of membrane-associated transport proteins. *Faseb J* 3: 2583–2592.
- Juvvadi SR, Glavy JS, Horwitz SB and Orr GA (1997) Domain organization of murine *mdr1b* P-glycoprotein: the cytoplasmic linker region is a potential dimerization domain. *Biochem Biophys Res Comm* 230: 442–447.
- Kamijo K, Kamijo T, Ueno I, Osumi T, Osumi T and Hashimoto T (1992) Nucleotide sequence of the human 70 kDa peroxisomal membrane protein: a member of ATP-binding cassette transporters. *Biochim Biophys Acta* 1129: 323–327.
- Kartner N, Evernden-Porelle D, Bradley G and Ling V (1985) Detection of P-glycoprotein in multidrug-resistant cell lines by monoclonal antibodies. *Nature* 316: 820–823.
- Kast C, Canfield V, Levenson R and Gros P (1996) Transmembrane organization of mouse P-glycoprotein determined by epitope insertion and immunofluorescence. *J Biol Chem* 271: 9240–9248.
- Katzmann DJ, Hallstrom TC, Voet M, Wysock W, Golin J, Volckaert G and Moye-Rowley WS (1995) Expression of an ATP-binding cassette transporter-encoding gene (YOR1) is required for oligomycin resistance in *Saccharomyces cerevisiae*. *Mol Cell Biol* 15: 6875–6883.
- Kelly A, Powis S, Kerr L, Mockridge I, Elliott T, Bastin J, Uchanska-Ziegler B, Ziegler A, Trowsdale J and Townsend A (1992) Assembly and function of the two ABC transporter proteins encoded in the human major histocompatibility complex. *Nature* 355: 641–644.
- Kool M, de Haas M, Scheffer GL, Scheper RJ, van Eijk MJ, Juijn JA, Baas F and Borst P (1997) Analysis of expression of cMOAT (MRP2), MRP3, MRP4, and MRP5, homologues of the multidrug resistance-associated protein gene (MRP1), in human cancer cell lines. *Cancer Res* 57: 3537–3547.
- Krogstad DJ, Gluzman IY, Kyle DE, Oduola AM, Martin SK, Milhous WK and Schlesinger PH (1987) Efflux of chloroquine from *Plasmodium falciparum*: mechanism of chloroquine resistance. *Science* 238: 1283–1285.
- Kroll JS, Hopkins I and Moxon ER (1988) Capsule loss in *H. influenzae* type b occurs by recombination-mediated disruption of a gene essential for polysaccharide export. *Cell* 53: 347–358.
- Kuchler K, Sterne R and Thorner J (1989) *Saccharomyces cerevisiae* STE6 gene product: a novel pathway for protein export in eukaryotic cells. *Embo J* 13: 3973–3984.
- Kuchler K and Thorner J (1992) Functional expression of human *mdr1* in the yeast *Saccharomyces cerevisiae*. *Proc Natl Acad Sci* 89: 2302–2306.
- Kupinski J, Coussen F, Bakalyar HA, Tang W, Feinstein PG, Orth K, Slaughter C, Reed R and Gilman A (1989) Adenylyl cyclase amino acid sequence: possible channel-or transporter-like structure. *Science* 244: 1558–1564.
- Lautier D, Canitrot Y, Deeley RG and Cole SPC (1996) Multidrug resistance mediated by the multidrug resistance protein (MRP) gene. *Biochem Pharm* 52: 967–977.
- Lemontt JF, Azzaria M and Gros P (1988) Increased *mdr* gene expression and decreased drug accumulation in multidrug-resistant human melanoma cells. *Cancer Res* 48: 6348–6353.
- Li ZS, Szczypka M, Lu YP, Thiele DJ and Rea PA (1996) The yeast cadmium factor protein (YCF1) is a vacuolar glutathione S-conjugate pump. *J Biol Chem* 271: 6509–6517.
- Lincke CR, Broeks A, The I, Plasterk RH and Borst P (1993) The expression of two P-glycoprotein (*pgp*) genes in transgenic *Caenorhabditis elegans* is confined to intestinal cells. *Embo Journal* 12: 1615–1620.
- Lincke CR, Smit JJ, van der Blik AM, van der Velde Koerts T and Borst P (1991) Structure of the human *MDR3* gene and physical mapping of the human *MDR* locus. *J Biol Chem* 266: 5303–5310.
- Lincke CR, The I, van GM and Borst P (1992) The P-glycoprotein gene family of *Caenorhabditis elegans*. Cloning and characterization of genomic and complementary DNA sequences. *Journal of Molecular Biology* 228: 701–711.
- Mahe Y, Lemoine Y and Kuchler K (1996) The ATP binding cassette transporters Pdr5 and Snq2 of *Saccharomyces cerevisiae* can mediate transport of steroids in vivo. *J Biol Chem* 271: 25167–25172.
- Malone PC (1990) A consensus structure for membrane transport. *Res Micro* 141: 374–383.
- Marger MD and Saier MH (1993) A major superfamily of transmembrane facilitators that catalyze uniport, symport and antiport. *Trend Bio Sci* 18: 13–20.
- Martin SK, Oduola AM and Milhous WK (1987) Reversal of chloroquine resistance in *Plasmodium falciparum* by verapamil. *Science* 235: 899–901.
- Mayer U, Wagenaar E, Beijnen JH, Smit JW, Meijer DKF, Vanasparen J, Borst P and Schinkel AH (1996) Substantial excretion of digoxin via the intestinal mucosa and prevention of long-term digoxin accumulation in the brain by the *mdr1a* P-glycoprotein. *British Journal of Pharmacology* 119: 1038–1044.
- McGrath JP and Varshavsky A (1989) The yeast STE6 gene encodes a homologue of the mammalian multidrug resistance P-glycoprotein. *Nature* 340: 400–404.
- Mosser J, Douar A, Sarde C, Kioschis P, Feil R, Moser H, Poustka A, Mandel J and Aubourg P (1993) Putative X-linked adrenoleukodystrophy gene shares unexpected homology with ABC transporters. *Nature* 361: 726–730.
- Naito M and Tsuruo T (1992) Functionally active homodimer of P-glycoprotein in multidrug resistant tumor cells. *Biochem Biophys Res Comm* 185: 284–290.
- Ng WF, Sarangi F, Zastawny RL, Veinot DL and Ling V (1989)

- Identification of members of the P-glycoprotein multigene family. *Mol Cell Biol* 9: 1224–32.
- Nielsen D and Skovsgaard (1992) P-glycoprotein as multidrug transporter: a critical review of current multidrug resistant cell lines. *Biochim Biophys Acta* 1139: 169–183.
- Nishi K, Yoshida M, Nishimura M, Nishikawa M, Nishiyama M, Horinouchi S and Beppu T (1992) A leptomycin B resistance gene of *Schizosaccharomyces pombe* encodes a protein similar to the mammalian P-glycoproteins. *Molecular Microbiology* 6: 761–769.
- Noda M, Ikeda T, Kayano T, Suzuki H, Takeshima H, Kurasaki M, Takahashi H and Numa S (1986) Existence of distinct sodium channel messenger RNAs in rat brain. *Nature* 320: 188–192.
- Nourani A, Wesolowski-Louvel M, Delaveau T, Jacq C and Delahodde A (1997) Multiple-drug-resistance phenomenon in the yeast *Saccharomyces cerevisiae*: involvement of two hexose transporters. *Mol Cell Biol* 17: 5453–60.
- O'Hare K, Murphy C, Levis R and Rubin G (1984) DNA sequence of the *white* locus of *Drosophila melanogaster*. *J Mol Biol* 180: 437–455.
- Ortiz DF, Keppel L, Speiser DM, Scheel G, McDonald G and Ow DW (1992) Heavy metal tolerance in the fission yeast requires an ATP-binding-cassette-type vacuolar membrane transporter. *Embo J* 11: 3491–3499.
- Ouellette M and Borst P (1995) New mechanisms of drug resistance in parasitic protozoa. *Ann Rev Micro* 49: 427–460.
- Ouellette M, Fase-Fowler F and Borst P (1990) The amplified H circle of methotrexate-resistant *Leishmania tarentolae* contains a novel P-glycoprotein gene. *Embo J* 9: 1027–1033.
- Ouellette M, Hettema E, Wust D, F. F-F and Borst P (1991) Direct and inverted DNA repeats associated with P-glycoprotein gene amplification in drug resistant *Leishmania*. *Embo J* 10: 1009–1016.
- Paulusma CC, Bosma PJ, Zaman GJR, Bakker CTM, Otter M, Scheffer GL, Scheper RJ, Borst P and Elferink R (1996) Congenital Jaundice In Rats With A Mutation In A Multidrug Resistance-Associated Protein Gene. *Science* 271: 1126–1128.
- Paulusma CC, Kool M, Bosma PJ, Scheffer GL, ter Borg F, Scheper RJ, Tytgat GN, Borst P, Baas F and Oude Elferink RP (1997) A mutation in the human canalicular multispecific organic anion transporter gene causes the Dubin-Johnson syndrome. *Hepatology* 25: 1539–1542.
- Price R, Robinson G, Brockman A, Cowman A and Krishna S (1997) Assessment of PFMDR 1 gene copy number by tandem competitive polymerase chain reaction. *Molecular and Biochemical Parasitology* 85: 161–169.
- Qin S, Xie A, Bonato MC and McLaughlin CS (1990) Sequence analysis of the translational elongation factor 3 from *Saccharomyces cerevisiae*. *J Biol Chem* 265: 1903–1912.
- Raymond M and Gros P (1989) Mammalian multidrug-resistance gene: correlation of exon organization with structural domains and duplication of an ancestral gene. *Proc Natl Acad Sci* 86: 6488–6492.
- Raymond M, Gros P, Whiteway M and Thomas DY (1992) Functional complementation of yeast *ste6* by a mammalian multidrug resistance *mdr* gene. *Science* 256: 232–234.
- Raymond M, Rose E, Housman D and Gros P (1990) Physical mapping, amplification and overexpression of the mouse *mdr* gene family in multidrug resistant cells. *Mol Cell Biol* 10: 1642–1651.
- Riordan J, Rommens J, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic M, Chou J, Drumm M, Iannuzzi M, Collins F and Tsui L (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 245: 1066–1073.
- Riordan JR, Deuchars K, Kartner N, Alon N, Trent J and Ling V (1985) Amplification of P-glycoprotein genes in multidrug-resistant mammalian cell lines. *Nature* 316: 817–819.
- Romsicki Y and Sharom FJ (1997) Interaction of P-glycoprotein with defined phospholipid bilayers: a differential scanning calorimetric study. *Biochemistry* 36: 9807–9815.
- Roninson IB, Ed. (1991) *Molecular and Cellular Biology of Multidrug Resistance in Tumor Cells*. New York, Plenum Press.
- Roninson IB, Abelson H, Housman DE, Howell N and Varshavsky A (1984) Amplification of specific DNA sequences correlates with multidrug resistance in Chinese hamster cells. *Nature* 309: 626–628.
- Roninson IB, Chin JE, Choi KG, Gros P, Housman DE, Fojo A, Shen DW, Gottesman M and Pastan I (1986) Isolation of human *mdr* DNA sequences amplified in multidrug-resistant KB carcinoma cells. *Proc Natl Acad Sci* 83: 4538–42.
- Ross JI, Eady EA, Cove JH, Cunliffe WJ, Baumberg S and Wootton JC (1990) Inducible erythromycin resistance in staphylococci is encoded by member of the ATP-binding transport super-gene family. *Mol Microbiol* 4: 1207–1214.
- Ruetz S, Delling U, Brault M, Schurr E and Gros P (1996) The *pfmdr1* gene of *Plasmodium falciparum* confers cellular resistance to antimalarial drugs in yeast cells. *Proc Natl Acad Sci* 93: 9942–9947.
- Ruetz S and Gros P (1994) Phosphatidylcholine translocase: a physiological role for the *mdr2* gene. *Cell* 77: 1071–1081.
- Samuelson J, Ayala P, Orozco E and Wirth D (1990) Emetine-resistant mutants of *Entamoeba histolytica* overexpress mRNAs for multidrug resistance. *Mol Biochem Para* 38: 281–290.
- Sanglard D, Ischer F, Monod M and Bille J (1996) Susceptibilities of *Candida albicans* multidrug transporter mutants to various antifungal agents and other metabolic inhibitors. *Antimicrob Agents Chemo* 40: 2300–2305.
- Sanglard D, Kuchler K, Ischer F, Pagani JL, Monod M and Bille J (1995) Mechanisms of resistance to azole antifungal agents in *Candida albicans* isolates from AIDS patients involve specific multidrug transporters. *Antimicrob Agents Chemo* 39: 2378–2386.
- Sarkadi B, Price EM, Boucher RC, Germann UA and Scarborough GA (1992) Expression of the human multidrug resistance cDNA in insect cells generates a high activity drug-stimulated membrane ATPase. *J Biol Chem* 267: 4854–4858.
- Savary S, Denizot F, Luciani M, Mattei M and Chimini G (1996) Molecular cloning of a mammalian ABC transporter homologous to *Drosophila white* gene. *Mammalian Genome* 7: 673–676.
- Schinkel AH, Mayer U, Wagenaar E, Mol CA, van Deemter L, Smit JJ, van der Valk MA, Voordouw AC, Spits H, van Tellingen O, Zijlmans JM, Fibbe WE and Borst P (1997) Normal viability and altered pharmacokinetics in mice lacking *mdr1*-type (drug-transporting) P-glycoproteins. *Proc Natl Acad Sci* 94: 4028–4033.
- Schinkel AH, Smit JJ, van Tellingen O, Beijnen JH, Wagenaar E, van Deemter L, Mol CA, van der Valk MA, Robanus-Maandag EC, te Riele HP, Berns AJM and Borst P (1994) Disruption of the mouse *mdr1a* P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. *Cell* 77: 491–502.
- Schinkel AH, Wagenaar E, van Deemter L, Mol CA and Borst P (1995) Absence of the *mdr1a* P-Glycoprotein in mice affects tissue distribution and pharmacokinetics of dexamethasone, digoxin, and cyclosporin A. *Journal of Clinical Investigation* 96: 1698–705.
- Schnell N, Engelke G, Augustin J, Rowenstein R, Ungermann V,

- Goetz F and Entian KD (1992) Analysis of the genes involved in the biosynthesis of lantibiotic epidermin. *Eur J Biochem* 204.
- Schoner B, Geistlich M, Rostock P, Rao RN, Seno E, Reynolds P, Cox K, Burgett S and Hersberger C (1992) Sequence similarity between macrolide resistance determinants and ATP-binding transport proteins. *Gene* 115: 93–96.
- Scotto KW, Biedler JL and Melera PW (1986) Amplification and expression of genes associated with multidrug resistance in mammalian cells. *Science* 232: 751–755.
- Sehested M, Simpson D, Skovsgaard T and Buhl-Jensen P (1989) Freeze-fracture study of plasma membranes in wild type and daunorubicin-resistant Ehrlich ascites tumor and P388 leukemia cells. *Virchows Arch [b]* 56: 327–335.
- Shaulsky G, Kuspa A and Loomis WF (1995) A multidrug resistance transporter/serine protease gene is required for pre-stalk specialization in *Dictyostelium*. *Genes and Development* 9: 1111–1122.
- Shen DW, Fojo A, Chin JE, Roninson IB, Richert N, Pastan I and Gottesman Mm (1986) Human multidrug-resistant cell lines: increased *mdr1* expression can precede gene amplification. *Science* 232: 643–645.
- Skach WR, Calayag MC and Lingappa VR (1993) Evidence for an alternate model of human P-glycoprotein structure and biogenesis. *J Biol Chem* 268: 6903–6908.
- Skovsgaard T (1978) Mechanism of cross-resistance between vincristine and daunorubicin in Ehrlich ascites tumor cells. *Cancer Res* 38: 4722–4727.
- Smit JJ, Baas F, Hoogendijk JE, Jansen GH, van der Valk MA, Schinkel AH, Berns AJ, Acton D, Nooter K, Burger H, Smith SJ and Borst P (1996) Peripheral neuropathy in mice transgenic for a human MDR3 P-glycoprotein mini-gene. *Journal of Neuroscience* 16: 6386–6393.
- Smit JJ, Schinkel AH, Oude Elferink RP, Groen AK, Waganaaf E, van Deemter L, Mol CA, Ottenhofer R, van der Lugt NM, van Room MA, van der Valk MA, Offerhaus GJ, Berns AJ and Borst P (1993) Homozygous disruption of the murine *mdr2* P-glycoprotein gene leads to a complete absence of phospholipid from bile and to liver disease. *Cell* 75: 451–462.
- Sonveaux N, Shapiro AB, Goormaghtigh E, Ling V and Ruyschaert JM (1996) Secondary and tertiary structure changes of reconstituted P-glycoprotein. A Fourier transform attenuated total reflection infrared spectroscopy analysis. *J Biol Chem* 271: 24617–24624.
- Sparreboom A, van Asperen J, Mayer U, Schinkel AH, Smit JW, Meijer DKF, Borst P, Nooijen WJ, Beijnen JH and Vantellingen O (1997) Limited oral bioavailability and active epithelial excretion of paclitaxel (taxol) caused by P-glycoprotein in the intestine. *Proceedings of the National Academy of Sciences of the United States of America* 94: 2031–2035.
- Sugawara I, Kataoka I, Morishita Y, Hamada H, Tsuruo T, Itoyama S and Mori S (1988) Tissue distribution of P-glycoprotein encoded by a multidrug-resistant gene as revealed by a monoclonal antibody, MRK 16. *Cancer Res* 48: 1926–1929.
- Szczyepka MS, Wemmie JA, Moye-Rowley WS and Thiele DJ (1994) A yeast metal resistance protein similar to human cystic fibrosis transmembrane conductance regulator (CFTR) and multidrug resistance-associated protein. *J Biol Chem* 269: 22853–22857.
- Tanabe T, Takeshima H, Mikami A, Flockerzi V, Takahashi R, Kanagawa K, Kojima M, Matsuo H, Hirose T and Numata S (1987) Primary structure of the receptor for calcium channel blockers from skeletal muscle. *Nature* 328: 313–318.
- Thiebaut F, Tsuruo T, Hamada H, Gottesman Mm, Pastan I and Willingham MC (1987) Cellular localization of the multidrug resistance gene product P-glycoprotein in normal human tissues. *Proc Natl Acad Sci* 84: 7735–7738.
- Thiebaut F, Tsuruo T, Hamada H, Gottesman Mm, Pastan I and Willingham MC (1989) Immunohistochemical localization in normal tissues of different epitopes in the multidrug transport protein P170: evidence for localization in brain capillaries and crossreactivity of one antibody with a muscle protein. *J Histochem Cytochem* 37: 159–164.
- Thomas PM, Cote GJ, Wohlk N, Haddad B, Mathew PM, Rabl W, Aguilar-Bryan L, Gagel RF and Bryan J (1995) Mutations in the sulfonylurea receptor gene in familial persistent hyperinsulinemic hypoglycemia of infancy. *Science* 268: 426–429.
- Tommasini R, Evers R, Vogt E, Mornet C, Zaman GJ, Schinkel AH, Borst P and Martinoia E (1996) The human multidrug resistance-associated protein functionally complements the yeast cadmium resistance factor 1. *Proc Natl Acad Sci* 93: 6743–6748.
- Townsend A and Trowsdale J (1993) The transporters associated with antigen presentation. *Seminars in Cell Biology* 4: 53–61.
- Treize AEO, Ratcliff R, Hawkins TE, Evans MJ, Freeman TC, Romano PR, Higgins CF and Colledge WH (1997) Co-ordinate regulation of the cystic fibrosis and multidrug resistance genes in cystic fibrosis knockout mice. *Hum Mol Genet* 6: 527–537.
- Treize AEO, Romano PR, Gill DR, Hyde SC, Sepulveda FV, Buchwald M and Higgins CF (1992) The multidrug resistance and cystic fibrosis genes have complementary epithelial expression. *EMBO J* 11: 4291–4303.
- Triglia T, Foote SJ, Kemp DJ and Cowman AF (1991) Amplification of the multidrug resistance gene *pfmdr1* in *Plasmodium falciparum* has arisen as multiple independent events. *Mol Cell Biol* 11: 5244–50.
- Tusnady GE, Bakos E, Varadi A and Sarkadi B (1997) Membrane topology distinguishes a subfamily of the ATP-binding cassette (ABC) transporters. *FEBS Letters* 402: 1–3.
- Ueda K, Cardarelli C, Gottesman Mm and Pastan I (1987) Expression of a full-length cDNA for the human 'MDR1' gene confers resistance to colchicine, doxorubicin, and vinblastine. *Proc Natl Acad Sci* 84: 3004–3008.
- Ueda K, Cornwell Mm, Gottesman Mm, Pastan I, Roninson IB, Ling V and Riordan JR (1986) The *mdr1* gene, responsible for multidrug-resistance, codes for P-glycoprotein. *Biochem Biophys Res Commun* 141: 956–962.
- Van Asperen J, Schinkel AH, Beijnen JH, Nooijen WJ, Borst P and van Tellingen O (1996) Altered pharmacokinetics of vinblastine in *Mdr1a* P-glycoprotein-deficient Mice. *Journal of the National Cancer Institute* 88: 994–9.
- Van der Bliek A, Van der Velde-Koerts T, Ling V and Borst P (1986) Overexpression and amplification of five genes in a multidrug resistant chinese hamster ovary cell line. *Mol Cell Biol* 6: 1671–1678.
- Van der Bliek AM, Baas F, Ten Houte de Lange T, Kooiman PM, Van der Velde Koerts T and Borst P (1987) The human *mdr3* gene encodes a novel P-glycoprotein homologue and gives rise to alternatively spliced mRNAs in liver. *Embo J* 6: 3325–3331.
- Van der Bliek AM, Baas F, Van der Velde Koerts T, Biedler JL, Meyers MB, Ozols RF, Hamilton TC, Joenje H and Borst P (1988) Genes amplified and overexpressed in human multidrug-resistant cell lines. *Cancer Res* 48: 5927–5932.
- Van der Bliek AM, Kooiman P, Schneider C and Borst P (1988) Sequence of *mdr3* cDNA encoding a human P-glycoprotein. *Gene* 71: 401–411.
- Van Helvoort A, Smith A, Sprong H, Fritzsche I, Schinkel A, Borst P and van Meer G (1996) MDR1 P-glycoprotein is a lipid translocase of broad specificity, while P-glycoprotein specifically translocates phosphatidylcholine. *Cell* 87: 507–517.

- Van Veen, HW, Callaghan R, Soceneantu L, Sardini A, Konings WN and Higgins CF (1998) A bacterial antibiotic-resistance gene that complements the human multidrug-resistance P-glycoprotein gene. *Nature* 391: 291–295.
- Volkman SK, Cowman AF and Wirth DF (1995) Functional complementation of the STE6 gene of *Saccharomyces Cerevisiae* with the Pfmdr1 gene of *Plasmodium Falciparum*. *Proc Natl Acad Sci* 92: 8921–8925.
- Walker JE, Saraste M, Runswick MJ and Gay NJ (1982) Distantly related sequences in the a and b subunits of ATP synthetase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *Embo J* 1: 945–951.
- Wei LY, Stutts MJ, Hoffman Mm and Roepe PD (1995) Overexpression of the cystic fibrosis transmembrane conductance regulator in NIH 3T3 cells lowers membrane potential and intracellular pH and confers a multidrug resistance phenotype. *Biophys J* 69: 883–895.
- Wellems TE, Panton LJ, Gluzman IY, do RVE, Gwadz RW, Walker JA and Krogstad DJ (1990) Chloroquine resistance not linked to mdr-like genes in a *Plasmodium falciparum* cross. *Nature* 345: 253–255.
- White TC (1997) Increased mRNA levels of ERG16, CDR, and MDR1 correlate with increases in azole resistance in *Candida albicans* isolates from a patient infected with human immunodeficiency virus. *Antimicrob Agents Chemo* 41: 1482–7.
- Wilson C, Seranno A, Wasley B, Bogenschutz M, Shankar A and Wirth D (1989) Amplification of a gene related to mammalian *mdr* genes in drug resistant *Plasmodium falciparum*. *Science* 244: 1184–1186.
- Wilson C, Volkman SK, Thaithong S, Martin RK, Kyle DE, Milhous WK and Wirth D (1993) Amplification of *pfmdr1* associated with mefloquine and halofantrine resistance in *Plasmodium falciparum* from Thailand. *Mol Biochem Para* 57: 151–160.
- Wu C-t, Budding M, Griffin M and Croop J (1991) Isolation and characterization of *Drosophila* multidrug resistance gene homologues. *Mol Cell Biol* 11: 3940–3948.
- Wu Y, Sifri CD, Lei HH, Su XZ and Wellems TE (1995) Transfection of *Plasmodium falciparum* within human red blood cells. *Proc Natl Acad Sci USA* 92: 973–977.
- Zhang JT, Lee CH, Duthie M and Ling V (1995) Topological determinants of internal transmembrane segments in p-glycoprotein sequences. *J Biol Chem* 270.
- Zhang JT and Ling V (1995) Involvement of cytoplasmic factors regulating the membrane orientation of P-glycoprotein sequences. *Biochem* 34: 9159–9165.

Address for correspondence: James M. Croop, MD, PhD Section of Pediatric Hematology/Oncology James Whitcomb Riley Hospital for Children, Room 2720, 702 Barnhill Drive, Indianapolis, IN 46202, U.S.A.