


### 2.3.5 Bilirubin metabolism

Namita Roy-Chowdhury, Yang Lu and Jayanta Roy-Chowdhury

**Bilirubin in medical history**

Bilirubin has attracted the attention of physicians since antiquity. Its chemistry, metabolism and disposal have been studied systematically during the last two centuries as a model for hepatic disposal of biologically important organic anions of limited aqueous solubility [1]. The discovery of several inherited disorders of bilirubin metabolism and excretion during the twentieth century has led to renewed interest in inherited diseases associated with jaundice, some of which continue to pose a therapeutic challenge, providing impetus for further research. While physicians are mainly concerned with the toxic effect of bilirubin and its importance as a liver function test, the antioxidant property of bilirubin may provide a physiological defence against oxidative injury.

**Formation of bilirubin**

**Sources of bilirubin**

Bilirubin is the breakdown product of the haem moiety of haemoglobin, other haemoproteins, such as cytochromes, catalase, peroxidase and tryptophan pyrrolysine, and a small pool of free haem. In humans, 250–400 mg of bilirubin is produced daily, of which approximately 20% is produced from non-haemoglobin sources [2]. Following the injection of radio-labelled porphyrin precursors (glycine or δ-aminolaevulinic acid), an ‘early-labelled peak’ of bilirubin (ELP) is excreted in bile within 72 h [3]. The initial component of ELP is derived mainly from hepatic haemoproteins. This is followed by a slower component, derived from both erythroid and non-erythroid sources, which becomes prominent in conditions associated with ineffective erythropoiesis, e.g. congenital dyserythropoietic anaemias, megaloblastic anaemias, iron-deficiency anaemia, erythropoietin-responsive transcription factor that regulates carbohydrate oxidation enzymes. *Mol Cell Biol* 20, 1868–1876.

**Enzymatic mechanism of bilirubin formation**

The microsomal haem oxygenase (HO) enzymes catalyse the oxidation of haem (Fig. 1). Three molecules of O₂ are consumed
in this reaction and a reducing agent, such as nicotinamide adenine dinucleotide phosphate hydrogenase (NADPH), is needed. The \( \alpha \)-methene bridge carbon is eliminated as CO, and the iron molecule is released \[6\]. Of the three forms of HO, HO-1 is ubiquitous and inducible by haem \[8\] and stress \[7\]; HO-2 is a constitutive protein, expressed mainly in the brain and the testis. The catalytic activity of HO-3 is low, and this protein may function mainly as a haem binding protein. CO produced by HO activity has a vasodilatory effect and regulates the vascular tone in the liver, heart and other organs during stress. Similarly, biliverdin and its product bilirubin are potent antioxidants, which may protect tissues under oxidative stress \[7,9\] (see below).

Biliverdin is reduced to bilirubin by the action of cytosolic biliverdin reductases, which require NADH or NADPH for activity \[10\]. As discussed later, bilirubin requires energy-consuming metabolic steps for excretion in bile. Thus, the physiological advantage of its formation is not clear. The strong antioxidant activity of bilirubin may be particularly important during the neonatal period, when other antioxidants are scarce in body fluids.

**Measurement of bilirubin production**

Bilirubin production can be quantified from the turnover of intravenously administered radioisotopically labelled bilirubin. Plasma bilirubin clearance is proportional to the reciprocal of the area under the radiobilirubin disappearance curve \[11\]. Bilirubin removal is calculated as the product of plasma bilirubin concentration and clearance. At a steady state of plasma bilirubin concentration, bilirubin removal equals bilirubin production. More conveniently, bilirubin formation can be quantified from CO, which is generated in equimolar amounts with bilirubin. Following rebreathing in a closed system, CO production is calculated from the CO concentration in the rebreathing mask and/or the increment in blood carboxyhaemoglobin saturation \[12\]. A small fraction of the CO may be formed by intestinal bacteria, which can be a significant source of CO in intestinal bacterial overgrowth syndromes \[13\].

**Inhibition of bilirubin production**

Substances, such as tin-protoporphyrin and tin-mesoporphyrin, that bind irreversibly to HO, but are not broken down, serve as ‘dead-end’ inhibitors of the enzyme and reduce bilirubin production \[14\]. Injection of tin-mesoporphyrin lowers serum bilirubin levels by 76% in neonates \[15\].

**Chemical characteristics of bilirubin**

The tetrapyrrole structure of bilirubin IX\( \alpha \) \((1,8\)-dioxo-1,3,6,7-tetramethyl-2,8-divinylbiladiene-a,c-dipropionic acid \[17\]) was solved by Fischer and Plieninger \[18\]. X-ray crystallography has revealed that the propionic acid side-chains of bilirubin form hydrogen bonds with the pyrrolic and lactam sites on the opposite half of the molecule, giving rise to a distorted ‘ridge tile’ structure \[19\] (Fig. 2). Formation of hydrogen bonds requires the interpyrrolic bridges at the 5 and 15 position of bilirubin to be in *trans* or ‘Z’ configuration, whereby bilirubin is termed bilirubin IX\( \alpha \)-ZZ. Engagement of all polar groups (two propionic acid carboxyls, four NH groups and two lactam oxygens) of bilirubin by the hydrogen bonds makes the molecule insoluble.
in water, necessitating chemical modification for excretion in bile. Disruption of the hydrogen bonds is accomplished \textit{in vivo} by enzyme-catalysed esterification of the propionic acid side-chains of bilirubin with contralateral NH groups and the lactam oxygen, thereby engaging all polar groups of the molecule and making it insoluble in water. Upon exposure to light, configurational changes (2 to E) occur at the C4 and C15 interpyrrolic bridges, disrupting the hydrogen bonds. The bilirubin $\text{IX}_{a}\text{-4E,15Z}$ configurational isomer can be cyclized forming the so-called lumirubin. These configurational and geometric isomers are more polar than the hydrogen-bonded bilirubin $\text{IX}_{a}\text{-4Z,15Z}$ and are excreted in bile without requiring glucuronidation.

The hydrogen bonds ‘bury’ the central methane bridge, so that the unconjugated bilirubin reacts very slowly with diazo reagents, whereas bilirubin glucuronides, which lack hydrogen bonds, react rapidly (‘direct’ van den Bergh reaction). The addition of ‘accelerators’ such as methanol, ethanol, 6 M urea or dimethyl sulphoxide to plasma disrupts the hydrogen bonds of bilirubin, so that both conjugated and unconjugated bilirubin react rapidly with diazo reagents (‘total’ van den Bergh reaction).

Bilirubin glucuronides in normal bile are 1-O-acyl conjugates linked to the propionic acid carboxyl of bilirubin in a $\beta$-O-ester linkage, which is hydrolysable by $\beta$-glucuronidase. However, during cholestasis, the migration of the 1-O-acyl bond from the C1 position to the C2, C3 or C4 position results in the generation of $\beta$-glucuronidase-resistant pigments [20], which are detectable in serum and bile by chromatographic analysis [21].
In cases of prolonged accumulation of conjugated bilirubin in plasma, as in cases of cholestasis or Dubin–Johnson syndrome, the pigment may become covalently bound to albumin [22]. This irreversibly protein-bound form, often termed delta-bilirubin, is included in the ‘direct’ fraction of bilirubin and is not eliminated in the bile or urine, which results in delayed clearance even after biliary obstruction or cholestasis is resolved.

**Effect of light**

The main absorption band of unconjugated bilirubin IXα is at 450–474 nm in most organic solvents. Upon exposure to light, the ‘E’ (trans) configuration of the 5 and/or 15 carbon bridges of bilirubin switches to the ‘E’ (cis) configuration. The resulting configurational isomers, ZE, EZ or EE, lack internal hydrogen bonds, are more polar than bilirubin IXα-EE and can be excreted in bile without conjugation [23]. The non-hydrogen-bonded molecule can be stabilized slowly by cyclization of the vinyl substituent in the endovinyl half of bilirubin IXα-EZ with the methyl substituent on the internal pyrrole ring, forming the stable structural isomer, E-cyclobilirubin. Because of its stability, this molecule is quantitatively important during phototherapy for neonatal jaundice [24]. Light and oxygen can also degrade a fraction of the bilirubin molecules into colourless fragments and biliverdin [25].

**Quantification of bilirubin**

Bile pigments can be quantified as native or derivatized tetrapyrroles, or after conversion to azoderivatives. Conversion to azopyrroles by reaction with diazo reagents is the most common method of measuring serum bilirubin levels in clinical laboratories. Electrophilic attack on the central bridge splits bilirubin into two diazotized azopyrrole molecules. As discussed above, conjugated bilirubin reacts rapidly (‘direct’ fraction), while total bilirubin is determined after adding an accelerator. Unconjugated bilirubin is calculated by subtracting the direct fraction from total bilirubin. As 10–15% of unconjugated bilirubin may give a ‘direct’ diazo reaction, this method slightly overestimates conjugated bilirubin.

Bilirubin and its conjugates in serum or bile can be quantified more accurately as intact bilirubin tetrapyrroles by high-pressure liquid chromatography [26–28]. Bilirubin mono- and diconjugates can be converted to methyl esters by alkaline methanolysis prior to separation [29] but, because the sugar groups are cleaved off, this method does not permit identification of specific conjugates.

For repeated bilirubin measurements in jaundiced infants, as an extension of clinical evaluation of jaundice, bilirubin levels can be assessed by measurement of the intensity of yellow discoloration of the skin using a special reflectance photometer [30]. Two slide tests (Ektachem) are available for determination of total bilirubin and the unconjugated, conjugated and irreversibly protein-bound fractions.

**Bilirubin toxicity**

Unconjugated bilirubin is toxic to many cell types, intracellular organelles and physiological processes. Bilirubin inhibits DNA synthesis [31] and ATPase activity of brain mitochondria [32], and uncouples oxidative phosphorylation. It has been reported to inhibit Ca2+-activated, phospholipid-dependent protein kinase C activity and cAMP-dependent protein kinase activity [33]. Which of these toxic effects is the predominant cause of bilirubin encephalopathy remains unclear at this time. Clinically, toxic effects of bilirubin, particularly on the brain, are seen in neonates and patients with severe inherited deficiency of bilirubin conjugation. Yellow discoloration of the hippocampus, basal ganglia and nuclei of the cerebellum and brain stem, found in infants with acute bilirubin encephalopathy, is termed kernicterus. Such discoloration is not found in patients with chronic encephalopathy, in whom focal necrosis of neurons and glia is seen [34].

As all toxic effects of bilirubin are abrogated by tight binding to albumin, cerebral toxicity is usually seen when there is a molar excess of bilirubin over albumin in plasma. At serum unconjugated bilirubin concentrations over 20 mg/dL, newborn babies are at risk of kernicterus. However, kernicterus can occur at lower concentrations in the presence of substances such as sulphonamides, radiographic contrast dyes and coumarin, which inhibit albumin–bilirubin binding by competitive or allosteric displacement [35,36]. Although immaturity of the blood–brain barrier in neonates has been implicated in the increased susceptibility of neonates to kernicterus, evidence to support this concept is insufficient. Normally, bilirubin entering the brain is cleared rapidly, but the pigment may bind to damaged and oedematous brain inhibiting its clearance, thereby increasing the susceptibility to bilirubin encephalopathy [37].

**Potential beneficial effects of products of haem breakdown**

Although clinicians are mainly concerned with the importance of bilirubin levels as a marker of liver disease and with the toxic effects of the pigment, biliverdin and bilirubin may exert some beneficial effects by virtue of their strong antioxidant properties. This may be relevant during the newborn period, when the level of other natural antioxidants is low. Bilirubin, which is toxic to neuronal cells at high concentrations, has been reported to have cytoprotective activity at lower concentrations. An inverse relationship between serum bilirubin levels and risk of ischaemic coronary artery disease has been observed [38], although whether such a protective effect extends to subjects with Gilbert syndrome is questionable [39]. Study of a large number of subjects in the United States has shown that the odds ratio for colorectal cancer is reduced to 0.295 in men and 0.186 in women per 1 mg/dL increment in serum bilirubin levels [40]. Similarly, a previous large study showed an inverse relationship between serum bilirubin levels and cancer mortality in a Belgian
population [41]. However, such associations do not conclusively prove a causative role for bilirubin, because possible confounding variables may exist.

**Bilirubin in body fluids**

About 4% of bilirubin in normal plasma is conjugated, but the clinical diazo-based methods overexpress this fraction (see above). In haemolytic jaundice, there is a proportional increase in plasma-unconjugated and -conjugated bilirubin. In contrast, in inherited disorders of bilirubin conjugation, the conjugated bilirubin is absent or reduced in proportion. In biliary obstruction or hepatocellular diseases, both conjugated and unconjugated bilirubin accumulate in plasma. Bilirubin is present in exudates and other albumin-containing body fluids and binds to the elastic tissue of skin and sclera. Haem in subcutaneous haematomas is sequentially converted to biliverdin and bilirubin, resulting in a transition from green to yellow discoloration. Because of tight binding to albumin, unconjugated bilirubin is not excreted in urine in the absence of albuminuria, but conjugated bilirubin, which is less strongly bound to albumin, appears in urine. Bilirubin is present in normal human bile predominantly as diglucuronide, with bilirubin monoglucuronide and unconjugated bilirubin accounting for less than 10% and 1–4% of the pigments respectively. In the presence of reduced bilirubin glucuronidating capacity of the liver, as in Gilbert syndrome and Crigler–Najjar syndrome type 2 (see Chapter 16.6), the proportion of bilirubin monoglucuronide increases to 30% or above. In addition to the glucuronides, small amounts of glucosyl, xylosyl and mixed conjugates of bilirubin are found in human bile.

**Disposition of bilirubin**

Disposition of bilirubin by hepatocytes comprises several specific steps, including transport of bilirubin to hepatocytes from sites of production, uptake by and storage within hepatocytes, enzyme-catalysed conjugation with glucuronic acid, active transport into the bile canalculus and degradation in the intestinal tract. These steps are summarized in Figure 3 and discussed briefly below.

---

**Fig. 3** Glucuronidation disrupts internal hydrogen bonding of bilirubin. Glucuronidation of the propionic acid carboxyl groups results in disruption of the internal hydrogen bonds, making the molecule more polar and secretable in bile. Disruption of hydrogen bonding exposes the central CHH bridge to diazo reagents, whereby bilirubin glucuronides give the direct van den Bergh reaction.
Transport in plasma

Unconjugated bilirubin circulates in plasma bound tightly but reversibly to albumin, which prevents its excretion in urine, except during albuminuria. Albumin binding keeps bilirubin in solution and abrogates its toxic effects. Conjugated bilirubin is bound less tightly to albumin, and the unbound fraction is excreted in the urine. As mentioned above, during prolonged conjugated hyperbilirubinaemia, a fraction of conjugated bilirubin becomes irreversibly bound to albumin. This fraction, termed delta-bilirubin, is not excreted in the bile or urine and disappears slowly, reflecting the long half-life of albumin [22]. A small unbound fraction of unconjugated bilirubin is thought to be responsible for its toxicity [42]. Albumin has one high-affinity primary binding site for bilirubin. Additional sites are occupied when bilirubin is in molar excess. Normal plasma concentration of albumin (500–700 μmol/L) exceeds that of bilirubin (3–17 μmol/L). However, during exaggerated neonatal jaundice and in patients with Crigler–Najjar syndrome, the molar concentration of unconjugated bilirubin may exceed that of albumin. Hypoalbuminaemia resulting from inflammatory states, chronic malnutrition or liver disease may precipitate bilirubin toxicity. Sulphonamides, anti-inflammatory drugs, cholecytographic contrast media, fusidic acid, azapropazone, sodium caprylate and N-acetyl tryptophan displace bilirubin from albumin and increase the risk of kernicterus in jaundiced infants [43]. Binding of short-chain fatty acids to albumin causes conformational changes, decreasing bilirubin binding. Because of its pathophysiological importance, various methods have been devised to measure the unbound fraction of bilirubin and the reserve albumin binding capacity. These include ultrafiltration, ultracentrifugation, gel chromatography, affinity chromatography on albumin agarose polymers, dialysis and electrophoresis. Rapid degradation of unbound bilirubin by H₂O₂ and horseradish peroxidase has been used to distinguish it from the bound fraction.

Uptake by hepatocytes

At the sinusoidal surface of the hepatocyte (Fig. 4), bilirubin dissociates from albumin and is taken up by the hepatocyte by facilitated diffusion that requires inorganic anions, such as Cl⁻. The protein(s) involved in sinusoidal bilirubin uptake have not been identified. A member of the organic anion transport protein family, termed OATP2 (also termed SLC21A6), has been proposed as the sinusoidal bilirubin transporter [44], but its importance in bilirubin transport has been questioned [45].

Storage within the liver cell

After entering the hepatocyte, bilirubin binds to the major cytosolic proteins, glutathione-S-transferases (GSTs, formerly designated ligandin or Y-protein). The GST proteins, which constitute 5% of the liver cytosol, bind various drugs, hormones, organic anions [46], a cortisol metabolite [47] and azo-dye carcinogens [48]. Bilirubin is a ligand for GSTs, but not a substrate for glutathione transfer. Binding to GSTs reduces the efflux of bilirubin from hepatocytes, thereby increasing its net uptake (Fig. 4). GST binding inhibits non-specific diffusion of bilirubin into various subcellular compartments, thereby preventing specific organelar toxicity, such as inhibition of mitochondrial respiration by bilirubin that is seen in vitro [49].

Conjugation of bilirubin

Conversion of unconjugated bilirubin to bilirubin diglucuronide or monoglucuronide by esterification of both or one of the propionic acid carboxyl groups is critical for efficient biliary excretion of bilirubin (Fig. 4).

Bilirubin uridine diphosphoglucuronate glucuronosyltransferase

Bilirubin is one of the many endogenous and exogenous substrates whose conjugation with glucuronic acid is mediated by one or more isoform of uridine diphosphoglucuronate glucuronosyltransferase (UGTs). UGTs are enzymes concentrated in the endoplasmic reticulum (ER) and nuclear envelope of many cell types [50]. They catalyse the transfer of the glucuronic acid moiety of UDP-glucuronic acid to the aglycone substrates, forming polar and usually less reactive products. Bilirubin glucuronidation is catalysed predominantly by a single UGT isoform, UGT1A1 [51]. The UGT superfamily of genes comprises two major families, UGT1 and UGT2. Nine isoforms within the UGT1A subfamily are expressed from a series of exons clustered in a unique manner on chromosome 2 at the 2q37 region [61]. Four consecutive exons (exons 2–5) located at the 3’ end of the UGT1A locus are used in nine different mRNAs. These encode the identical carboxy-terminal domains of these UGT isoforms, which contain the UDP-glucuronic acid binding site. Upstream of these four common region exons is a series of unique exons, each preceded by a separate promoter. Only one of these exons is utilized in a specific UGT mRNA. The unique exon encodes the variable N-terminal domain of the nine different UGT isoforms that impart aglycone specificity to the individual isoforms. Depending on which promoter is used, transcripts of various lengths are generated. In all cases, the unique exon, located at the 5’ end of the transcript, is spliced to exon 2, and the intervening sequence is spliced out. The exons are named according to the unique first exon. Thus, UGT1A1 utilizes the unique exon 1A, UGT1A6 utilizes exon 1A6, etc. [53].

The presence of a separate promoter upstream from each unique region exon permits differential regulation of individual UGT isoforms during development and in response to inducing agents. UGT1A1 is expressed after birth [54] and is induced by phenobarbital and clofibrate [55]. Delayed expression of UGT1A1 is a major cause of neonatal hyperbilirubinaemia in primates. Treatment of rats with triiodothyronine markedly
reduces UGT activity towards bilirubin, whereas the activity towards 4-nitrophenol is increased [56].

In humans, the expression of UGT1A1 is limited to hepatocytes and, to a lesser extent, the proximal small intestine. UGTs are integral to ER membranes. In addition to the enzyme content, UGT1A1 activity is affected by the lipids of the ER membrane. UGT activity in native microsomal vesicles is latent [57], probably because the ER membranes pose a barrier to the polar sugar donor UDP-glucuronic acid or as a result of the constraint of the enzyme by the membranes. Based on hydrophobicity analysis, the major portion of mature UGT molecules, including the UDP-glucuronic acid and the aglycone binding sites, is thought to be located within the ER cisternae. There is a single 17-amino-acid membrane-spanning segment and a 26-amino-acid cytoplasmic tail at the carboxy-terminal end of the molecule. Full enzyme activity is manifested in vitro by treatment of the microsomes with membrane-permeabilizing agents, such as digitonin or alamethacin. UDP-N-acetylgalactosamine (UDP-glucNac) stimulates the internalization of UDP-glucuronic acid into intact microsomal vesicles and is thought to be the natural activator of UGTs within hepatocytes. UGT1A1 forms homodimers within the ER membrane, which may be required for its full catalytic activity [58]. In addition, it may interact with other UGT isoforms, as well as other proteins of the ER.

**Canalicular excretion of conjugated bilirubin**

Conjugated bilirubin undergoes unidirectional transport into the bile against a concentration gradient, so that bilirubin concentration in the bile can be as high as 150-fold that in the hepatocyte. The electrochemical gradient of −35 mV, generated by the sodium pump, may help in the canalicular transport but, by itself, is too small to account for this large concentration gradient. The energy for the uphill transport of bilirubin and many other non-bile salt organic anions is derived from adenosine triphosphate (ATP) hydrolysis by the canalicular ATP-binding cassette protein, ABCC2 [also termed the MDR-related protein 2 (MRP2) or the multispecific organic anion transporter, MOAT]. ABCC2 pumps glutathione-, glucuronic acid- or sulphate-conjugated compounds across the canalicular membrane [59,60]. Canalicular transport of organic anions is unidirectional from the cytoplasm of the hepatocyte into the bile. Canalicular transport may be assisted by the membrane potential, but the contribution of membrane potential in organic anion transport has not been quantified. Mutant animals that lack ATP-dependent canalicular transport of non-bile acid organic anions retain normal activity with respect to potential-driven canalicular transport of non-bile acid organic anions, including bilirubin glucuronides [60]. The ATP-dependent canalicular organic anion transport is mediated by a canalicular membrane protein, termed canalicular multispecific organic anion transporter (cMOAT) or MRP2 [61].

Maximal bilirubin secretory capacity (Tmax) into the bile canaliculus depends on bile flow, which has bile salt-dependent and non-bile salt-dependent components. Bile acids increase the trafficking of vesicles containing MRP2 and the bile salt export pump (BSEP) from the Golgi apparatus to the apical domain of hepatocyte plasma membranes, thereby increasing the concentration of the transporters in the canalicular membrane [61].

**Fate of bilirubin in the gastrointestinal tract**

Although conjugated bilirubin is not substantially absorbed from the intestines, a fraction of the small amount of unconjugated bilirubin that is excreted in bile is absorbed and undergoes enterohepatic circulation. In situations in which increased amounts of unconjugated bilirubin are excreted in bile, such as e.g. during phototherapy for neonatal jaundice or Crigler-Najjar syndrome, absorption of unconjugated bilirubin from the intestine may be clinically significant [62]. In these cases,
interruption of bilirubin reabsorption by ingestion of various substances, including calcium salts, can enhance the effect of phototherapy [63].

Degradation of bilirubin by intestinal bacteria generates urobilinogen and related products [64]. A major portion of the urobilinogen reabsorbed from the intestine is excreted in bile, but a small fraction is excreted in urine. Urobilinogen is colourless; its oxidation product, urobilin, contributes to the colour of normal urine and stool. During severe intrahepatic cholestasis or complete obstruction of the bile duct, urobilinogen and urobilin are absent in urine and stool, resulting in pale (so-called clay-coloured) stool. In liver disease and states of increased bilirubin production, urinary urobilinogen excretion is increased.

**Alternative routes of bilirubin elimination**

In the absence of bilirubin glucuronidation, a fraction of bilirubin is excreted as hydroxylated products [65], probably by the action of microsomal P450s [66] and mitochondrial bilirubin oxidase in liver [67] and other tissues.

During intrahepatic or extrahepatic cholestasis, conjugated bilirubin accumulates in plasma. In total biliary obstruction, renal excretion becomes the major pathway of bilirubin excretion [68]. Renal excretion of conjugated bilirubin depends on glomerular filtration of the non-protein-bound fraction of conjugated bilirubin.

**References**

25. McDonagh AF (1975) Thermal and photochemical reactions of bilirubin IX. *Ann NY Acad Sci* 244, 553–569.


2.3.6 Metabolism of bile acids

Peter L.M. Jansen and Klaas Nico Faber

Introduction

Bile acids are synthesized in the liver from cholesterol; they are secreted in bile and stored in the gallbladder. After a meal, the gallbladder contracts, and stored bile is transferred to the duodenum and via the jejunum to the ileum. This movement is stimulated by intestinal propulsion. In the ileum, 90–95% of bile salts are reabsorbed and returned to the liver. The remainder is lost to the colon, where primary bile salts are transformed by bacterial metabolism into secondary bile salts. Some of the secondary bile salts are also reabsorbed, and the rest is removed with the faeces. Primary and secondary bile salts return to the liver via the portal circulation. In the liver, bile salts are taken up into hepatocytes, thereby completing the enterohepatic cycle.

Bile acids serve a number of functions: (i) they are the main solutes in bile and, as such, they are important for the generation of the so-called bile salt-dependent bile flow; (ii) bile salts are indispensable for the secretion of cholesterol and phospholipids from the liver; (iii) in bile, bile salts form mixed micelles that keep fat-soluble organic compounds in solution, including fat-soluble vitamins; (iv) in the intestine, bile salts promote the dissolution and hydrolysis of triglycerides by pancreatic enzymes; (v) bile salts act as signalling molecules in the regulation of enzymes and transporters of drug and intermediary metabolism.

The adult human liver produces about 500 mg of bile acids per day [1,2]. About three times this amount represents the total bile acid pool size that cycles through the enterohepatic circulation [2]. Bile acids complete an enterohepatic cycle about eight times per day. Enterohepatic cycling represents an efficient system for reusage of active components. Enterohepatic cycling not only serves to reclaim bile acids, but it also enables bile acids to act as messengers that carry signals from intestine to liver. Thus, they regulate their own synthesis and transport rates. Bile acids are also able to repress hepatic fatty acid and triglyceride synthesis [3,4].

Biosynthesis and metabolic defects

At least 16 different enzymes are involved in the biosynthesis of bile salts [1,5,6]. Most of these enzymes are active in the neutral (or classic) and acidic (or alternative) pathways, the two main routes for the conversion of cholesterol to the primary bile acids cholic acid (CA) and chenodeoxycholic acid (CDCA) (Fig. 1). The neutral pathway starts with the hydroxylation of the sterol nucleus of cholesterol by 7α-hydroxylase (CYP7A1) in the endoplasmic reticulum. CYP7A1 is regarded as the rate-limiting enzyme in bile acid biosynthesis, exemplified by the fact that mice deficient for Cyp7a1 have a 75% reduced bile acid pool size causing vitamin deficiencies, lipid malabsorption and liver failure [7–9]. The acidic pathway starts with the hydroxylation of the cholesterol side-chain by sterol 27-hydroxylase (CYP27). The CYP27 product, 5-cholesten-3β-27-diol, is not a substrate for CYP7A1, but is hydroxylated at the C7 position by an alternative P450 enzyme, CYP7B1. From here on, the neutral and acidic pathways largely overlap. Double hydroxylated CDCA and triple hydroxylated CA are the principal bile acids. Their ratio depends on the activity of sterol 12α-hydroxylase (CYP8B1). Bile acid synthesis is completed in hepatocyte peroxisomes, where bile acid coenzyme A:amino acid N-acyltransferase (BAAT) conjugates either taurine or glycine to CA or CDCA. At least 95% of the bile acid pool is generated through these two pathways. Extensive intracellular transport of bile acid intermediates occurs between various organelles. Transport in and out of these organelles may be mediated by transport proteins, but these have not been characterized in detail yet.

Bile acid synthesis defects (BASD) are rare genetic disorders that are the underlying cause of approximately 2% of persistent cholestasis in infants (see also Chapter 16.10, Genetic cholestatic diseases). BASDs are recognized by the absence or reduction of normal primary bile salts in serum and/or urine. Instead, non-typical bile acids and sterols are often detected in the body fluids of these patients. These can be identified by fast atom bombardment-ionization–mass spectrometry (FAB-MS) and gas chromatography–mass spectrometry (GC-MS). Disease-causing mutations have been identified in 9 out of the 16 bile acid biosynthesis enzymes (Table 1). Cholestasis is a common clinical presentation of these diseases. The associated liver diseases may vary from mild to life-threatening but, in many cases, can be managed by replacement of deficient primary bile salts. This not only leads to restoration of normal bile function, but also induces feedback inhibition on the production of toxic bile acid intermediates.

Patients with CYP7A1 deficiency have a markedly reduced bile acid synthesis rate [10]. Symptoms include hyperlipidaemia, premature vascular disease and gallstones. A mutation in the CYP7A gene that results in truncation of the enzyme has been detected in these patients. Only one case of CYP7B1 deficiency has been reported to date [11]. This child produced no primary bile acids, and serum concentrations of the toxic 27α-hydroxy cholesterol were increased. A mutation was identified in the CYP7B1 gene that truncates and inactivates the enzyme. In addition, it was found that expression of CYP7A, at both the mRNA and activity level, was absent. Bile acid treatment was ineffective, suggesting that the biosynthesis of toxic 27α-hydroxy cholesterol cannot be suppressed.