Pathogenesis of Gallstones

HARRY F. WEISBERG, M.D.

The Ida Soref-David and Ruth Coleman
Department of Pathology and Laboratory Medicine,
University of Wisconsin Medical School,
Milwaukee Clinical Campus,
Mount Sinai Medical Center,
Milwaukee, WI 53201

ABSTRACT

The three lipids in bile, cholesterol, lecithin, and bile salts (about 90 percent of the dry weight of normal gallbladder bile) are amphipathic substances having both hydrophobic and hydrophilic functional groups. Knowledge of the physicochemical factors of gallstone formation (especially cholesterol stones) has increased in the past two decades. The absolute amount of cholesterol supersaturation determines the extent of cholesterol precipitation. The ionic strength of the bile and the types of bile salts present are minor factors, whereas the ratios of bile salts to lecithin at a particular concentration of total lipids are the major factors contributing to gallstone production. Bile acids (salts) form micelles which allow the lecithin and cholesterol to dissolve within the micelles. Thus the administration of bile acids allows for non-invasive dissolution of some cholesterol gallstones. Additional important risk factors are genetic and ethnic, sex (females predominate), obesity, diet (in contrast to animal protein and more refined carbohydrate diets, there is less lithogenicity with diets containing plant protein and unrefined carbohydrates), certain diseases, and drug therapy. Pigment stones make up the majority of radiopaque stones and are predominant in the Orient; they are seen in certain diseases and in infections of the biliary tree.

Introduction

Gallstones have been recognized in man and animals since antiquity. The gallstones from a mummy of the priestess of Amenem (1500 B.C.) were lost when the mummy was destroyed during the bombing of London in World War II. Gallstones are common in herbivorous animals but rare in carnivera; stones have been found in both the elephant and horse, animals without a gallbladder. The stones from cattle served as a major source of yellow pigment used by the ancient artists. It is estimated that from 16 to 20 million people in the United States have gallstones. Annual cholecystectomies vary from 80,000 in Canada to about one third of a million in the United States. Annually, 5,000 to 8,000 deaths in the United States are attributed to gallstone disease.
In 1500, Paracelsus advanced the theory that certain chemical disturbances in the body initiated the precipitation of impurities in the biliary ducts. Various theories were promulgated since then. Gallstones are formed by the precipitation of insoluble constituents of bile. By tradition, gallstones have been classified as (1) pure cholesterol or pigment (largely made of calcium bilirubinate), (2) mixed—two or three components of cholesterol, calcium bilirubinate, and calcium carbonate (usually more than 70 percent cholesterol), or (3) combination stones with a nucleus of one type and a shell of another substance. It is very rare to have a chemically pure gallstone. Analysis by infrared spectroscopy and X-ray diffraction of gallstones found in the gallbladder and in the common bile duct revealed five major components—cholesterol, calcium bilirubinate, fatty acid calcium salts, inorganic calcium salts, and black pigment material. Gallbladder stones contained black pigment and inorganic calcium salts more frequently than did the common duct stones. However, the ductal stones had an incidence of 55.5 and 18.9 percent for calcium bilirubinate and fatty acid calcium salts, respectively, in contrast to 8.7 and 1.2 percent for the gallbladder stones.

Though relatively rare, the precipitation of calcium carbonate in the gallbladder, as a separate mass or upon already existing stones, may give rise to difficulties in the proper interpretation of a cholangiogram. This whitish precipitate has been termed Kalkmilchgalle or milk of calcium bile or limey bile. White bile, however, denotes the absence of bile in the bile ducts; this has also been referred to as acholia, the term for lack of bile in the intestinal tract.

**Cholesterol Stones**

The word cholesterol is derived from the Greek, chol (bile or gall) and steroids. Cholesterol gallstones usually contain more than 70 percent cholesterol and account for most of the gallstone disease seen in the Americas, Europe, and Africa. If such cholesterol stones contain enough calcium, they may be radiopaque (33 percent) but 80 to 86 percent of the radiolucent stones consist of cholesterol.

The solubility of cholesterol in bile fluid is limited and is found in the crystalline form in patients with cholesterol gallstones. Cholesterol is insoluble in water but normal human bile contains about 10 mmol of cholesterol dissolved in one liter of bile. Cholesterol is a C27, unsaturated, monohydroxylated, nonpolar compound; the hydroxyl group allows for surface solubility on water but no solubility in water. The solubility of cholesterol depends on the relative concentrations of the three major lipid components in bile: conjugated bile salts, phospholipids (at least 90 percent in the form of lecithin—phosphatidyl choline), and free cholesterol. These three lipids make up about 90 percent of the dry weight of normal gallbladder bile. The total solids in gallbladder bile can vary from 3 to 30 g per dl or 2.8 to 24.9 g per dl (average 12 g per dl). The concentration varies from patient to patient, from time to time in the same patient, and from one collection site to another. Hepatic bile is more dilute (1 to 4 g per dl) and the concentration in hepatic duct bile ranges from 0.2 to 7.9 g per dl (average 2.7 g per dl). The absolute solubility of cholesterol varies with the concentration of total solids; in dilute hepatic bile it is about three mole percent whereas it is as high as 10 mole percent in a very concentrated gallbladder bile. The vast majority of normal and abnormal bile samples reported in the literature have relatively limited ranges of concentration.

Free or unconjugated bile acids are the end product of cholesterol degradation (primary bile acids); they are C24.
saturated, mono- or poly-hydroxylated, polar compounds with the carboxyl group located in the side chain. The generic term, bile salt, refers to bile acids conjugated in the liver with taurine or glycine or to bile alcohols esterified with a sulfate group. This conventional description has been changed since naturally occurring bile acids are a more numerous and diverse group than has been recognized. Thus, C_{27} bile acids are found in the bile, serum, and urine of infants with a familial form of cholestasis and intrahepatic bile duct anomalies and in the bile and serum of patients with Zellweger syndrome. Patients with Zellweger syndrome also have a C_{29} dicarboxylic bile acid in their serum, and a C_{23} bile acid has been identified in the serum of an adult with cholestasis. The meconium of normal full-term infants has been shown to have several, short side-chain bile acids (C_{20}, C_{21}, and C_{22}).

The three lipid components of bile are amphipathic molecules, possessing groups with characteristically different properties, e.g., hydrophobic group at one end and hydrophilic group at the other end. The steroid nuclei of cholesterol and the bile acids and the two long-chained fatty acids of lecithin are hydrophobic and seek the oil phase. In contrast, the polar hydroxyl group(s) of cholesterol and the bile salts, the taurine and glycine conjugates of bile, and the phosphoryl choline and glycerol ester groups in lecithin are hydrophilic and are attracted to the water phase. Though lecithin is insoluble in water, water penetrates between the hydrophilic choline groups, resulting in a swelling of the lecithin—“liquid crystals.” The water-insoluble cholesterol can interdigitate between the lecithin molecules with its steroid nucleus buried in the fatty acid interior of the lecithin molecule. A “micelle” is a polymer of about 50 to 100 amphipathic molecules, e.g., bile salts, arranged spherically, usually with the hydrophobic end on the inside and the hydrophilic group on the outside. The aggregation to form simple micelles occurs above specific concentrations of the bile salts in the water (the “critical micellar concentration”). Lecithin can dissolve in these bile-salt micelles, resulting in mixed micelles in which the cholesterol can dissolve in the hydrophobic core between the fatty acid chains of the lecithin. Water associated with the polar groups of lecithin swells the micelles, allowing the incorporation of more cholesterol than could be contained by the simple bile-salt micelles alone.

In figure 1 is shown a triangular phase diagram for plotting relative percent of the total molar concentrations (millimoles per liter) of the three major components of bile. The original phase diagram was based on the average composition of normal bile—90 percent water and 10 percent solids (bile salts, cholesterol, and lecithin). With biles obtained at surgery or directly from the gallbladder or hepatic duct, one can determine the absolute concentration of solids and thus can estimate the true solubility of cholesterol. However, with bile samples obtained by duodenal drainage (cholecystokinin stimulation), the total bile lipids are diluted. Therefore one uses an arbitrary concentration (e.g., 10 percent total solids) to obtain a rough estimate of cholesterol solubility in duodenal bile. From a practical point of view, one should look for crystals in freshly collected bile; if present, they are separated by centrifugation and the remaining liquid phase analyzed. Centrifugation of crystals which have appeared some time after collection gives an erroneous analysis. Biles are often frozen (−20°C) for transport; this may alter the chemical and physical state of the bile.

The concentration of the lipid components is expressed in terms of “substance” concentration (millimoles per liter) rather than “mass” concentration (milligrams per deciliter). Bile has 3 to 23 mmol per l of cholesterol, 30 to 50
Figure 1. Triangular phase diagram showing physical state of combinations of cholesterol, lecithin, and bile salts. ZONE I: One phase; micellar liquid. ZONE II: Two phases; micellar liquid and cholesterol monohydrate crystals. ZONE III: Three phases; micellar liquid, cholesterol monohydrate crystals, and liquid crystals of lecithin and cholesterol. ZONE IV: Two phases; liquid crystals of lecithin and cholesterol and isotropic liquid. Modified from Carey and Small.4

mmol per l of lecithin, and 40 to 145 mmol per l of bile salts. Adding the lower values gives a total of 73 mmol with corresponding data of 4, 41, and 55 percent of total lipids for cholesterol, lecithin, and bile salts, respectively. Adding the upper values gives a total of 218 mmol with the corresponding values of 10, 23, and 67 percent of total lipids for cholesterol, lecithin, and bile salts.

Figure 1 is based on 20 g per dl solutions of total lipids (cholesterol, lecithin, and sodium taurocholate) in 0.15 molar sodium chloride at room temperature.4 Studies with X-ray analysis and polarizing microscopy resulted in the description of the different number and types of equilibrium phases. The equilibrium phases for Zones II, III, and IV are found with concentrations of lecithin of less than 20, between 20 to 40, and above 40 percent of total moles, respectively.4

In figure 2 is shown an enlargement of the clinically applicable portion of the phase diagram. It shows point P at a concentration of cholesterol at 5, bile salts at 80, and lecithin at 15 percent of total moles; the specimen is not saturated with cholesterol and lies in Zone I, a single phase of micellar liquid. Line ABC is the maximal effective solubility of cholesterol in varying mixtures of lecithin and bile salts. The line AB is too high, and the true limit of solubility of cholesterol at 37°C is shown by line DBC. Mixtures of bile with the value plot above line ABC contain readily precipitable excess cholesterol. Mixtures falling in the metastable area (ABDA) have a slight excess of cholesterol; no precipitation occurs un-
less the mixture is seeded or nucleated or it stands for long periods.

In 1954 Isakkson\(^7\) reported that bile from patients with cholesterol gallstones had a ratio of cholesterol to the sum of bile salts and phospholipids greater than 1:11 whereas most patients without gallstones had a lower ratio. This was confirmed and extended by Admirand and Small\(^1\), who reported the percent cholesterol saturation.\(^{13}\) Similar expressions are the lithogenic index\(^{11}\) and the cholesterol saturation index.\(^{17}\) The indices are expressed as a fraction of unity instead of percentage. The percent saturation is determined by extending a line from point P to the apex of the triangular diagram (figure 2); the intersection (X) of this line with line ABC gives the relative cholesterol concentration at 100 percent effective saturation. Point X is at about 10 percent cholesterol; thus, the original relative concentration of 5 percent (point P) is divided by 100 to give 50 percent saturation; if expressed as the lithogenic index, the result is 0.50. A normal bile is less than 100 percent saturated with cholesterol or has a lithogenic index less than 1.00; an index value above 1.00 denotes a saturated (more than 100 percent cholesterol saturation) or lithogenic bile.\(^{11}\)

In contrast to drawing a line to the apex of the triangular coordinates (figures 1 and 2), Thomas and Hofmann\(^{19}\) modified the calculations of the lithogenic index by utilizing rectangular coordinates; these are easily adapted to computerized calculations. The relative percent concentration of cholesterol is plotted on the ordinate and the abscissa depicts the ratio of the “percent” lecithin to the sum of the lecithin and bile salt percentages. They utilized the data of Admirand and Small\(^1\) and used regression analysis to obtain a third-degree polynomial; the equation for figure 3 is

\[
y = 4.86 + 39.3x - 74.4x^2 + 0.88x^3.
\]

The original data for point P (figure 2) had concentrations of bile salts, lecithin, and cholesterol at 80, 15, and 5 percent, respectively, of the total moles. Thus the ratio of 15/15 + 80 = 0.158. In figure 3, the vertical line at 0.158 gives a cholesterol concentration of about 9.5 at point X; thus 5/9.5 = 0.53 for the lithogenic index or 53 percent cholesterol saturation.

Carey and Small\(^4\) criticize this for-

---

**Figure 2.** Triangular coordinate plot of maximal effective cholesterol solubility line (ABC) and true equilibrium solubility line (DBC). C, cholesterol; L, lecithin; and BS, bile salts. Area within ABDA represents metastable region. After Small.\(^1,15,16\)
mula, stating it is not as accurate as their fifth-degree polynomial expression; however, they have utilized that same rectangular plot of cholesterol on the ordinate versus the ratio of lecithin to the sum of lecithin and bile salts on the abscissa. They showed that, within physiological bile salt:lecithin ratios at 37°C, the influence of type of bile salt and ionic strength is minor whereas the effects of bile salt:lecithin ratio and the total lipid concentration are major factors. Thus, utilizing cholesterol saturation values appropriate to the total lipid concentration, all cholesterol stone patients have supersaturated gallbladder biles,—132 percent for normal weight and 199 percent for morbidly obese individuals. For control and pigment stone patients, the mean values were 95 and 98 percent, respectively, even though about half of the biles were supersaturated. Cholesterol crystals were seen in 83 percent in gallbladder and 58 percent in hepatic bile of cholesterol stone patients but were not seen in controls or pigment stone patients.4 Fasting biles of normal individuals may be supersaturated with cholesterol.

Stages

Cholesterol gallstone disease can be described in five stages.16 Stage I involves the genetic, biochemical, or metabolic defect leading to supersaturation; this can be subdivided into six types. Type 1 is the excessive loss of bile salts as seen in ileal disease or surgery or congenital loss of ileal active transport of bile salts. The decreased reabsorption leads to a decreased bile salt pool and bile salt secretion rate. Type 2 is an oversensitive bile acid feedback seen most usually in stone disease present in non-obese Caucasians. There is a relative depression of bile acid synthesis; the decreased hepatic return excessively inhibits bile acid synthesis. Type 3 exhibits excessive cholesterol secretion despite a normal bile salt secretion rate. This is noted in obese patients with increased synthesis of cholesterol; the bile acid pool may be in the normal range. Type 4 is a mixture of Types 2 and 3 and is seen in the native American Indians and, perhaps, many Caucasians.

The last two types are primarily extrahepatic in origin. Type 5 exhibits a rapid bile salt circulation seen in patients with a decreased bile acid pool, the small amount of normal bile in the gallbladder cannot compensate for abnormal bile entering the gallbladder during fasting. Type 6 includes disorders of the gallbladder, ducts, or sphincters; aseptic or bacterial cholecystitis may secondarily complicate other types of gallstone disease, including pigment stone disease.16

Stage II is the chemical phase wherein
the bile is supersaturated with cholesterol but stones are absent. Studies of duodenal drainage reveal excess cholesterol (figure 1) but no crystals are seen on microscopic examination of the fluid. The physical stage (III) shows cholesterol crystals on microscopy (nucleation, flocculation, and precipitation) but no stones are evident by cholecystography. Stones are present in Stages IV (growth) and V (clinical). In Stage IV macroscopic stones or a non-functioning gallbladder is seen on cholecystography. The duodenal drainage study usually reveals cholesterol crystals or abnormal bile, but the patient is asymptomatic. In the clinical stage (V), signs and symptoms are present with the stones causing blockage of the cystic duct, cholecystitis, and jaundice.16

**Risk Factors**

*Genetic and Ethnic.* The incidence of cholesterol gallstones varies from country to country and is related to the extent of cholesterol saturation of the bile (lithogenic index). The extremes are represented by the Masai tribe of East Africa who, despite a high fat diet and high cholesterol absorption with bile saturation of about 50 percent, do not have cholesterol stones to the young Arizona Pima Indian women with bile saturation of about 110 percent and 80 percent prevalence of stones. Swedish individuals revealed about 60 percent and 150 percent bile saturation in 1954 and 1971 studies, respectively, but both groups had a prevalence of gallstones at 50 percent.13 Younger female siblings of women with gallstones have more saturated bile than similar siblings of control patients without gallstones.

*Age.* Gallstones have been reported from fetus to extreme old age but the average patient is in the fifth decade at the time of diagnosis or surgery. The incidence increases with age.

*Sex.* The frequency of cholelithiasis is from two to four times greater in women and occurs at an earlier age in women. Pregnancy has been implicated in the higher incidence; with pregnancy there is an increase of the lithogenic index, ascribed to the effect of estrogen (estriol). Exogenous estrogens (for contraception and post-menopausal replacement) increases the incidence of gallstones. In addition, in the last trimester of pregnancy, gallbladder emptying is impaired and hypercholesterolemia is evident.

*Obesity.* In obese individuals there is an increased biliary secretion of cholesterol secondary to increased cholesterol synthesis. During active weight reduction, bile saturation increases as a result of decreased secretion of all biliary lipids; in some, secretion of bile acids decreases more than that of cholesterol. Once the weight is stabilized, the bile acid pool expands to normal but cholesterol secretion remains low with resulting decreased cholesterol saturation.

*Diet.* Urbanization and adoption of western dietary habits has changed the composition of gallstones in Japan,—cholesterol stones are increasing and the pigment stones are decreasing. A higher caloric intake has been associated with individuals with gallstones and with the cholesterol concentration of T-tube bile in post-cholecystectomy patients. Other dietary factors that have been suggested are high cholesterol, polyunsaturated fats, high carbohydrate, and low vegetable fiber.

Animal protein (casein) is more lithogenic and cholesterolemic than plant protein (soy protein); a diet of equal amounts of animal and vegetable protein is only slightly more cholesterolemic than a diet containing only plant protein.9 The addition of bran to the diet of patients with gallstones had reduced the lithogenic index of the bile; the reports are conflicting when bran is added to the
diet of normal individuals. The type of carbohydrate consumed has been shown to affect the bile cholesterol saturation. Subjects with probable gallstones, given a refined carbohydrate diet (refined sugar, white flour, and white rice) had an increased saturation index (1.5 ± 0.10) whereas an unrefined carbohydrate diet (only whole grain products) resulted in an index of 1.2 ± 0.12. Long term parenteral nutrition results in the formation of biliary sludge and gallstones. Confirmation of these observations is reported in an animal model (prairie dog).

**Diseases.** Malabsorption of bile acids from the ileum disturbs the enterohepatic circulation, diminishing the bile acid pool and the rate of secretion of bile; this results in the bile becoming lithogenic. This is seen in ileal disease or ileal resection, Crohn’s disease of the small bowel, and cystic fibrosis with pancreatic insufficiency. The higher incidence of gallstones in diabetics of both sexes is related to the obesity seen in some diabetics and the associated increase of cholesterol secretion in the bile. Patients with hyperlipoproteinemia, especially Type IV (hypertriglyceridemia), have a high incidence of cholesterol gallstones.

**Drugs.** Clofibrate (ethyl chlorophenoxyisobutyrate), used to treat severe types of hyperlipidemia, lowers plasma cholesterol by inhibiting cholesterol synthesis. It mobilizes cholesterol stores out of the tissues with increased secretion into the bile and also impairs bile salt synthesis; the net result is lithogenic bile with gallstone formation. The same may hold true for nicotinic acid.

**Pigment Stones**

Pigment gallstones result from abnormal metabolism of bile pigment and are composed of bile pigment, calcium, and a matrix of organic material. The cholesterol content has been described as none, less than 5 percent, and as “trace amounts (less than 20 percent).” Half of the pigment stones are opaque to X-rays, but two-thirds of the radiopaque stones are of the pigment variety. A small pigment stone is found in the center of almost every “mixed” cholesterol gallstone. The center consists of calcium bilirubinate as a protein-pigment complex, often containing copper, mixed with calcium phosphate and calcium carbonate. Even the “pure” cholesterol stone may have some pigment in its center. Pigment stones are predominant in the Orient; however, 5 to 15 percent, 10 to 12 percent, or 27 to 33 percent of gallstones in Occidentals are of this variety.

Pigment gallstones are more common with increasing age but do not appear to be related to sex or obesity. Certain disease states are associated with an increased incidence of pigment stones: cirrhosis of the liver, particularly alcoholic, and in severe chronic hemolytic anemias seen in sickle cell disease, thalassemia major, congenital spherocytosis, erythroblastosis fetalis or resulting from aortic valve replacement. Biliary infection, especially with *E. coli*, or parasitic infections (e.g., ascaris) that cause stasis (and then infection) lead to deconjugation of the soluble bilirubin diglucuronide, resulting in the insoluble free bilirubin and glucuronic acid. The deconjugation is catalyzed by β-glucuronidase. The biliary tree is normally free of this lysozymal enzyme but coliform infections can elaborate the enzyme or the enzyme may be produced by the gallbladder epithelium. Stasis per se can initiate nonenzymatic hydrolysis. The normal biliary tree contains glucaro-1, 4-lactone (D-glutaric acid) which inhibits the β-glucuronidase. In the presence of sufficient calcium, the free bilirubin is precipitated as calcium bilirubinate or similar compounds in the biliary tract. These compounds then polymerize to form a stone. Thus, the pigment stones
may be bile pigment-calcium stones or pure pigment stones (non-bilirubin pigments with calcium, copper, or iron). \textsuperscript{14}

References


