Summer 2021

Studies on protein solubility

Ke Li

Department: Chemical and Biochemical Engineering

STUDIES ON PROTEIN SOLUBILITY

by

KE LI

A DISSERTATION

Presented to the Graduate Faculty of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

CHEMICAL ENGINEERING

2021

Approved by:

Dr. Daniel Forciniti, Advisor
Dr. Douglas Ludlow
Dr. Xinhua Liang
Dr. Fateme Rezaei
Dr. Risheng Wang

PUBLICATION DISSERTATION OPTION

This dissertation consists of the following three articles, formatted in the style used by the Missouri University of Science and Technology:

Paper I, found on pages 12-41, is submitted to Process Biochemistry Journal.

Paper II, found on pages 42-97, has been published in the Journal of Chemical & Engineering Data

Paper III, found on pages 98-135, is intended for submission to Scientific Report.

ABSTRACT

Solubility is a critical factor of protein-based drugs during processing and patient administration. This study focused on two aspects of solubility: one was the poly (ethylene) glycol (PEG) precipitation of monoclonal antibodies (mAbs) and another was crystallins aggregation that is associated with cataracts. Protein precipitation by PEG is a common technique for downstream processing. The effects of pH, ionic strength and the exclude volume effect on the protein precipitation by PEG were extensively studied, but the effects of glycosylation on protein precipitation by PEG has not been examined. Protein aggregation is not only a problem in downstream processing, but it is also related to some diseases. In 2015, lanosterol and 25-hydroxycholesterol were reported to dissolve protein aggregates in cataract lenses. Researchers focused on dissolution of lens protein aggregates, but the effects of those two sterols on the formation of aggregates were not investigated. The objectives of this dissertation were: 1) to determine the role of glycosylation in the precipitation of mAbs by PEG, and 2) to study the effects of lanosterol and 25-hydroxycholesterol on α-crystallin aggregation. The glycosylated mAbs showed higher solubility than non-glycosylated mAbs. It was found that available solubility models cannot correlate the effect of glycosylation. Lanosterol and 25-hydroxycholesterol failed to prevent α-crystallin aggregation. A concentration of 125μM of the two sterols promoted the aggregation of α-crystallin and of the α/γ complex possible by serving as nucleation sites. The secondary and tertiary structures of α-crystallin were not affected upon addition of the two sterols. The α-crystallin chaperon activity and the capacity of binding with Cu^{2+} were not affected either.

ACKNOWLEDGMENTS

First, I would like to express my sincere gratitude to my advisor, Dr. Daniel Forciniti for his valuable guidance and encouragement through deepest waters when most doors seemed closed. Thanks to him, I have grown both professionally and personally learning what it takes to balance life.

I would also like to express my deepest gratitude to Dr. Douglas Ludlow, Dr. Fateme Rezaei, Dr. Xinhua Liang and Dr. Risheng Wang for their time and endeavor in serving as committee members and reviewing this dissertation. A special thanks to all the staff in the Department of Chemical and Biochemical Engineering for their help and support during my research.

Thanks to my parents for their prayers and believing and supporting me throughout the journey. Special thanks to my friends for their wonderful support. Finally, I would like to acknowledge everybody who stood by me through thick and thin especially during this COVID-19 pandemic and helped me reach the other side.

I would like to dedicate this dissertation to my parents Yafei Li and Yafei Wei for their unconditional love and support throughout the journey.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

PAPER I Page

PAPER II

NOMENCLATURE

Symbol	Description
GlcNAc	N-Acetylglucosamine
MALS	2.3 Multiple Angle Laser Light Scattering
UV	Ultra Violet
ε	extinction coefficients
GC-MS	gas chromatography-mass spectrometry
DMSO	dimethyl sulfoxide
NRTL	Non-random two-liquid model
UNIQUAC	Universal quasichemical
DSC	Differential Scanning Calorimeter
DIOP	diisooctylphthalate

1. INTRODUCTION

With the development of recombinant technology, the commercial production of medicinal protein preparations became a reality. The preparation of a protein as a pharmaceutical drug became an indispensable part of the pharmaceutical industry. However, the solubility of proteins during processing and storage affected its medical and commercial applications. Proteins are the final product of DNA translation. Proteins serve as enzymes, carriers of cell signaling and ligand binding, and are structural components. Each protein has distinctive structures that support their biological functions. The protein solubility is a determining factor for their biological functions. For example, keratin is a type of insoluble protein that is the structural material for hair and nails[1]; crystallins require high solubility to maintain eye lens transparency.[2] Protein aggregation is a consequence of soluble protein molecules forming insoluble aggregate. Undesired protein aggregates lead to diseases like cataracts.

In this dissertation, protein solubility was studied from two different perspectives. One was the poly (ethylene) glycol (PEG) precipitation of monoclonal antibodies (mAbs) and the other was the effects of lanosterol and 25-hydoxycholesterol on the α-crystallin aggregation processes.

1.1. PROTEIN STUCTURE

Proteins have four levels of organization that makes each protein type unique. The primary structure of a protein refers to the amino acid sequence that forms the peptide chain; i.e. the sequence of amino acid residues in the protein molecule, including their

number, type and sequence. Amino acids are connected to each other by peptide bonds, i.e. a molecule of water is removed between the α-amino group of one amino acid and the α-carboxyl group of another amino acid. Peptide bonds have the properties of partial double bonds, so the entire peptide unit is a rigid planar structure. The end of the polypeptide chain containing a free amino group is called the N-terminus of the peptide chain, and the end containing a free carboxyl group at the other end is called the C-terminus of the peptide chain. Changes in the primary structure of a protein can change its secondary structure and its function.[3]

Secondary structure refers to the structure formed by the folding of the polypeptide chain backbone. The most basic types of secondary structure are α-helical and β-sheet structures; both structures are maintained by hydrogen bonds. There are also β-turns and random coils. Each peptide bond in the helix is involved in the formation of hydrogen bonds to maintain the stability of the spiral. β-sheet structures are also common. In this structure, the polypeptide chain exists in a relatively stretched form, and the arrangement of the peptide chains (or peptide segments) can be parallel or antiparallel. The axial distance between amino acids is 0.35 nm, and adjacent peptide chains are connected to each other by hydrogen bonds to form a layered structure.[4]

The tertiary structure of a protein is the three-dimensional shape of the entire polypeptide chain formed by further folding and rolling of secondary structures. Amino acids interact with one another via charge-charge, hydrophobic, disulfide, or other interactions. The polypeptide chains of rhe protein are coiled and folded in multiple directions in a three-dimensional space to form a tight, approximately spherical structure. The space inside the molecule can only accommodate a few water molecules. Almost all

polar side chains are distributed on the outer surface of the molecule to form a hydrophilic shell, while most non-polar side chains are buried inside the molecule and do not contact water. The interaction of side chain in protein molecules plays an important role in stabilizing the tertiary structure of proteins.[5]

The quaternary structure refers to the structure of a protein formed by interactions between multiple polypeptide chains. In a protein with a quaternary structure, each peptide chain with a tertiary structure is called a subunit. The absence of a subunit makes the protein biologically inactive. The quaternary structure involves the spatial arrangement of subunits in the entire molecule and the relationship between subunits.[6]

1.2. PROTEIN SOLUBILITY

The solubility of a protein is affected by changes in pH, temperature, ionic strength, the addition of cosolvents[7], and by post-translational modifications like glycosylation. Each of these environmental factors directly affects protein solubility. The pH and ionic strength both affect the net charge of protein. A few amino acids are weak acids and basis; therefore, proteins normally bear a net charge. At the pH equal to the isoelectric point (pI) of a protein, the protein molecules exist in the form of zwitterions, and their net molecular charge is zero (that is, the positive and negative charges are equal). At pHs below or above the pI, the protein is positive-charged or negative-charged respectively. For a pH near the pI, protein's solubility is at a minimum because the electrostatic repulsive force was minimized.[8] When the force is weakened, intermolecular collisions lead to aggregates that eventually precipitate. Therefore, when the pH of a protein solution is at the pI, the protein solubility is the lowest and a

precipitate is most likely to form. Many physical properties at the pI, such as viscosity, swelling, and osmotic pressure are reduced, which is beneficial to the filtration of the suspension.[9] The effects of ionic strength on the protein solubility depend on pH. Briefly, at a pH near pI the addition of salt to a protein solution first increases (salting-in) and then decreases (salting out) protein solubility. At pHs below or above pI, the addition of salt first decreases protein solubility by screening electrostatic repulsions and then it raises protein solubility due to 1) the interaction between weakly hydrated monovalent anions and polar and nonpolar groups on the protein surface; 2) the interaction between amide bond and multivalent cation.[10]

The solubility of a native protein also decreases as the temperature decreases. An increase in temperature unfolds the protein; therefore, it induces the formation of aggregates that eventually precipitate. High temperatures not only affect the secondary structure of a protein, but also, in some cases, it alters the quaternary structure of a protein oligomer, i.e. the α-crystallin that belong to the small heat shock protein form larger oligomers at temperatures higher than 37^{o}C.

Dehydration of a protein is the main mechanism for protein destabilization caused by the addition of cosolvents. Protein precipitation by alcohols has a long history. The alcohols disrupt the hydration shell of protein that leads to aggregation and precipitation. To prevent irreversible protein aggregation, alcohol precipitation is usually conducted at 4 ^{o}C.

Non-ionic polymers include dextran and poly (ethylene) glycol (PEG) were used to precipitate proteins. This phenomenon were explained by exclude volume theory that was first introduced by Asakura and Oosawa.[11] The non-ionic polymer molecules were

excluded from the volume between two protein molecules, then the volume between two protein molecules becomes a phase of pure solvent; therefore, the particular distribution of the polymer causes a pressure imbalance that pushes the proteins against each other, known as the exclude volume theory. PEG precipitation has been used, for example, for the crystallization of glycosylated and non-glycosylated variants of agglutinin, extraction of chicken IgY from egg yolk[12], precipitation of lysozyme[13], and coupled with chromatography to purify botulinum neurotoxin type B, among other uses.[14]

Post-translational modifications also affects protein solubility. Common post-translational modifications include: acylation, acetylation, alkylation, glycosylation *etc.*[15] Glycosylation consists of the addition of various glycan groups to the polypeptide chain.[16] Glycosylation may protect proteins from stresses that destabilize them, such as precipitants, pH, chemicals, and heat.[17] Aggregation is a known phenomenon observed in commercial antibodies preparations (particularly after reconstituting a freeze-dried sample).[18] The formation of biologically inactive aggregates decreases their efficacy. Moreover, there is evidence that protein aggregates present in a protein drug increase immunogenicity.[19, 20] It was argued that glycosylation increased the stability of the protein; therefore, it was expected that a non-glycosylated antibody was more susceptible to aggregation.[21]

Precipitation is an important downstream processing step. Precipitants are used to lower protein solubility and induce protein solid-liquid or liquid-liquid phase-separation to separate proteins from undesired impurities. Salts, organic solvents, and non-ionic polymers (PEG and dextran) are frequently used precipitants. One of the advantages to used PEG is that it precipitates proteins without. pH, ionic strength, and temperature are

critical factors in PEG precipitation,[22-24]and they were extensively studied as evident in previous related literature. Depletion and electrostatic forces control the extent of protein precipitation by PEG. However, the role of glycosylation is unclear, creating a gap in the literature.

In paper I, glycosylated and non-glycosylated mAbs were selected as models to study the effect of glycosylation on protein precipitation by PEG. PEG 1450 Da and 8000 Da were used. The effects of pH and temperature were also explored. Additional studies were performed in the presence of a Griffonia (Bandeiraea) simplicifolia Lectin II (GSL-II), which binds to glycosylated proteins. The precipitation curves were fitted with a Cohn salting-out equation analogous. The PEG precipitation efficiency coefficient, which was extracted from the curve fitting, was compared to the available solubility models. This study's data showed that glycosylation enhanced the mAbs solubility in the presence of PEG.

1.3. DISSOLUTION OF CATARACTS BY THE ADDITION OF STEROLS

Mammals' eye lenses have onion-like layered structures. The outward facing edge of the lens consists of a mono-layer of epithelial cells that differentiate to new fiber cells during the development of lens over the lifetime of an individual.[25] In order to maintain the transparency of the lens, the fiber cells lack of blood vessels and the sub-cellular structures. The lenses' high refractive index is caused by the high concentration of crystallins expressed in fiber cells, which require high solubility of crystallins. Conversely, lens fiber cells lack the capacity for protein turnover and repair.[26] The degradation of crystallins accumulate over a lifetime and thus their solubility decreases.

The low solubility of aged crystallins results in aggregation, and they finally increase the scattering of light and form cataracts.[27, 28]

Cataracts cause approximately 50% of blindness worldwide.[29] Cataract surgery is readily available in the developed world, but it is less common in underdeveloped countries. Surgery is invasive and requires relatively sophisticated equipment; additionally, well-trained physicians are scarce in the underdeveloped world. Moreover, surgeries are almost nonexistent in large underprivileged populations in Asia, Africa, and the Middle East as well as in Central and South America. Curing cataracts using eye drops is a very attractive and financially sound alternative.

Zhao *et al.*[30] discovered that lanosterol (a triterpenoid and a precursor of cholesterol that accumulates in the eye lenses) can dissolve congenital cataracts in rabbits and dogs. Their work was inspired by the observation that a population with congenital cataracts was deficient in the enzyme that participates in one of the synthesis steps of lanosterol. A year later, Shanmugam *et al.*[31] tested lanosterol for the solubilization of age-related cataracts in humans; however, they observed that the triterpenoid (used in the dosage and protocol used by Zhao *et al.* to treat congenital cataracts) was inefficient in the solubilization of senile cataracts. At the same time, Makley *et al.*[32] used differential scanning calorimetry to determine the effect of approximately 2,500 compounds on the melting temperature of the model heat shock protein Hsp27, which is similar to the αB-crystallin. A promising set of 32 sterols was checked for their binding capabilities to a mutant of the αB-crystallin (R120G αB-crystallin). This mutation of αB-crystallin was known to destabilize proteins in the lenses. Makley *et al.* also found that the most promising compound (25-hydroxylcholesterol) increased transparency in mouse model

cataracts. This 25-hydroxylcholesterol compound was bound to R120G αB-crystallin and lowered its melting temperature (the temperature at which the protein unfolds). They concluded that lanosterol was not a good candidate because of its low solubility and marginal effect on the melting temperature of Hsp27; therefore, it was not included in the final set of 32 sterols. These observations started a new chapter in cataract research.[33]

Following the studies of lanosterol and 25-hydroxycholesterol to restore lens' transparency, other researchers tried to repeat those experiments or collect additional experimental evidence about the activity of those sterols. Shen *et.al.*[34] found that, in vitro, lanosterol (40 μM in M199 medium) delayed the occurrence of lens opacity in a lanosterol synthase inhibited rat lens. Xu *et.al.*[35] used 20 μM lanosterol in 1% DMSO to successfully reverse W151R mutant human βB2-crystallin aggregates. Kang *et al.*[36] used all atom molecular dynamics simulation and free energy perturbation techniques to show that lanosterol can bind to the hydrophobic interface of dimers of human γD-crystallin preventing aggregation. Daszynski *et al.*[37] failed to repeat the experiments by Zhao *et al.*[30] and Markley *et al*[32] using their same approach. Also Daszynkski *et al.*'s docking simulations shows those two sterols cannot bind to the groove which is formed by the α-crystallin dimer using two wild types (PDB 2WJ7 and 2KLR) and a R120G mutant (PDB 2Y1Z) αB-crystallin. Nagai *et.al.*[38, 39] used lanosterol nanoparticles with a particle size distribution from 50 to 400 nm to repair the space and structural collapse in the early stages in the lenses. They found that it delays the onset of opacification of the lenses with a remarkable lens structure collapse and opacification, but it does not repair them. They speculated that the repeated injection of lanosterol

nanoparticles attenuated the manifestation of cataract-related factors and perhaps protects the lenses from oxidative stresses.

A plausible explanation for the restoration of transparency in the lenses is the fitting of lanosterol and 25-hydroxycholesterol into the interface of crystallin dimers. Makley *et al.*[32] and Daszynski *et al.*[37] studied the possible interactions between 25-hydroxycholesterol or lanosterol and the α-crystallin dimer interface by docking simulations. Both of them predicted the K_d (dissociation constant that is the ligand concentration at which half the protein molecules will have a ligand bound) of the two sterols were in the high micromolar or even millimolar range; i.e. K_d values for lanosterol-2KLR (wildtype αB-crystallin) and 25-hydroxycholesterol dimer interface were 73.63 mM 25-hydroxycholesterol and 1.22 mM, respectively.[37] Daszynski *et al.*[37] concluded that this high concentration of the two sterols could not be achieved clinically in the lens.

α-crystallin exists in the lenses as a multimeric aggregate.[40] The aggregates dissolve when stressed (α-crystallin is a member of the small heat shock protein's family) releasing dimeric αA-crystallins (αA-crystallin and αB-crystallin) that stabilize other proteins, such as β- and γ-crystallins,[41] in the lenses. Members of the sHsp are found in all forms of life and have a highly conserved α-crystallin domain structure across species. The sHsps are molecular chaperones that bind to unfolded or partially unfolded proteins, preventing their interaction with other unstable proteins. Unfortunately, the effect of sterols on the α-crystallin chaperone activity is not investigated.

Evidence of cataract lens restoration only occurred in animals (dogs and rabbits). The ineffectiveness in restoring human cataracts is the limitation of lanosterol solubility in the delivery medium. Mixed solvents showed the ability to enhance lanosterol solubility. The solubility of lanosterol is not reported in the open literature. A lanosterol product datasheet from Cayman Chemical Company[42] reported that the lanosterol solubility in ethanol and N,N-dimethylformamide (DMF) are 0.25mg/mL and 3 mg/mL, respectively, but the temperature was not specified. Zhao *et al.* used 10-40 μM of lanosterol in 1% DMSO that may contain undissolved lanosterol particles, which raises a question that the crystallins aggregates interact with dissolved lanosterol or the undissolved ones. . Access to solubility data for lanosterol is critical for its use to restore crystallin aggregates, and it is also important for patient administration, especially through eye drops.

In paper II, lanosterol solubility in organic solvent and water-alcohol binary systems was measured first. A lab-built experiment set-up was used to measure lanosterol solubility at different temperatures. It was found that lanosterol solubility increased with increased temperatures, and it increased with increasing alcohol content. Lanosterol has a low solubility in water at approximately 0.5 μM at 25^{o}C and increased little with increasing temperatures. The melting temperature and enthalpy of fusion for lanosterol was also measured. Powder X-Ray diffraction (PXRD) was employed to monitor the crystal form of the undissolved lanosterol. We demonstrated that aqueous solutions that contained more than 0.5 μM lanosterol had undissolved lanosterol particles.

After collecting the lanosterol solubility data, a series of experiments were performed using lanosterol,25-hydroxycholesterol and crystallins. Following the studies

of lanosterol and 25-hydroxycholesterol, other researchers tried to repeat those experiments and to study the mechanism of the restoration ability of the two sterols. The effects of lanosterol and 25-hydroxycholesterol on the α-crystallin aggregation process were not investigated in the available literature. If the compounds cannot inhibit protein aggregation, they cannot solubilize aggregates. Furthermore, the influence of the two sterols on the α-crystallin critical biological function and the chaperone activity were not covered in previous researchers' work.

In paper III, a series of experiments were conducted using bovine crystallins isolated from raw bovine lenses. The aggregation kinetics of α-crystallin incubated with lanosterol and 25-hydroxycholesterol with a concentration of 125 μM or lanosterol at 0.5 μM and of 25-hydroxycholesterol at 45 μM at 55°C were monitored by turbidity, high performance liquid chromatography (HPLC), and Fourier-transform infrared spectroscopy (FTIR). Then the chaperone activity of α-crystallin using γ-crystallin as substrates in the presence of those two sterols was evaluated using the same methods, but the temperature was lowered to 50°C. The α-crystallin binding capacities of Cu^{2+} were measured by the PAR colorimetric method and the bis-ANS fluorescence assay. Those two sterols failed to prevent α-crystallin aggregation and could not enhance α-crystallin chaperone activity regardless of concentration. FTIR and Circular dichroism (CD) results showed that the secondary and tertiary structures of α-crystallin were not affected by the sterols. Furthermore, the α-crystallin binding capacity of Cu^{2+} was not affected by those two sterols.

PAPER

I. POLY (ETHYLENE) GLYCOL (PEG) PRECIPITATION OF GLYCOSYLATED AND NON-GLYCOSYLATED MONOCLONAL ANTIBODIES

ABSTRACT

The solubility of monoclonal antibodies (mAb) affects their production and their intravenous administration to patients. In this work, the solubility of a fully glycosylated and a non-glycosylated human mAb expressed in corn was studied by inducing their precipitation by adding poly(ethylene) glycol (PEG). The experiments were done using PEG 1,450 and 8,000 at concentrations ranging from 0 to 30 % w/w, at different pHs and temperatures. Additional studies were performed in the presence of a Griffonia (Bandeiraea) simplicifolia Lectin II, which binds to glycosylated proteins. These studies clearly show that glycosylation increases the solubility of the antibody. These studies also show that models based on excluded volume principles or on the statistical correlation of solubilities are unable to capture the effect of glycosylation on protein precipitation by PEG.

1. INTRODUCTION

In the past decades, monoclonal antibody (mAb) and mAb fragment-based drugs have drawn great attention in the pharmaceutical industry because of their high target specificity and therapeutic efficacy.[1] mAbs are glycoproteins, the glycan groups help

conserving the structure and function of the antibody. [2] Also the different types of glycan groups make antibodies very heterogeneous. Moreover, the glycan groups affect the immunological properties of antibodies by altering their affinity for Fc receptors.[3] Antibodies glycosylated in vitro or in transgenic plants show greater diversity than their in vivo counterpart and this diversity may differ from batch to batch, which causes some batches not meeting specifications. A straightforward approach to eliminate this problem is to shut down the glycosylation machinery altogether. Although the heterogeneity problem is of course eliminated, the fact that the native and the non-glycosylated antibodies are not exactly the same (in spite of both having the same biological activity) raises questions about the solution stability (solubility) of the non-glycosylated antibody.[4, 5]

Glycosylation may protect proteins from stresses that destabilize them such as precipitants,[6-8] pH,[9, 10] chemicals,[11, 12] and heat.[10, 13] Destabilization of the protein may lead to aggregation, which is a known phenomenon observed in commercial antibodies preparations [14, 15] (particularly after reconstituting a freeze-dried sample). The formation of biologically inactive aggregates decreases the solubility of the preparation impeding their normal function. Moreover, there is evidence that protein aggregates present in a protein drug may increase immunogenicity.[16] Because it has been argued that glycosylation increases the stability of the protein, it is expected that a non-glycosylated antibody will be more susceptible to aggregation.[3]

Proteins may be precipitated by manipulating the pH, temperature, ionic strength, and the chemical composition of the solution.[17] Non-ionic polymers such as dextran and poly(ethylene)glycol (PEG) can also be used to precipitate proteins.[18, 19] PEG is

more effective than dextran[20] and the precipitation effectiveness increases with an increase in the size of PEG; however, solutions containing high molecular weight PEG (>10,000) are so viscous that they are not practical to use.[21]

PEG precipitation has been used, for example, for the crystallization of glycosylated and non-glycosylated variants of agglutinin,[22] extraction of chicken IgY from egg yolk,[23] precipitation of lysozyme,[24] and coupled with chromatography to purify botulinum neurotoxin type B[25] among many other uses. In addition, PEG precipitation has been used as a screening tool for developing high protein concentration formulations.[26] Recently, a new method of continuous PEG precipitation followed by tangential flow filtration was developed.[27] Such a method may be used to replace the costly chromatography method (mostly Protein A based) in the downstream processes of mAbs.

Two arguments have been used to explain the mechanism of protein precipitation by non-ionic-polymers. The simplest one consists of considering that the presence of the polymer decreases the volume of the solution available to the protein. This would force the proteins to be closer to each other such as precipitate nuclei would form. The second argument consists of considering that because of steric constraints the polymers are excluded from the space separating two protein molecules. This particular distribution of the polymer causes a "pressure imbalance" that pushes the proteins against each other. Purely stearic arguments (argument 1) fail to explain the effect of the protein charge on protein precipitation by non-ionic polymers. On the contrary, the second argument can nicely (and rigorously) incorporate the effect of pH on protein precipitation by non-ionic polymers.[18] Because both approaches assume that the solvent is a continuous fluid,

they cannot capture the contribution of hydration forces to protein stability. For example, water of hydration enters into Mahadevan-Hall's[28] model as a contribution to the volume of the molecule. The consequence of that is that the more hydrated the protein is the more prone to precipitation it will be, which does not agree with experimental evidence.

In this paper, the precipitation by the addition of PEG of a monoclonal antibody expressed in corn was studied. Glycosylated and non-glycosylated samples of the same antibody were used. The experimental parameters were PEG concentration and molecular weight, pH and temperature. A few experiments were done with the addition of Griffonia (Bandeiraea) simplicifolia Lectin II (GSL-II), which is a lectin[29] specific to the GlcNAc group that is present in the corn-expressed antibody. The solubility data were fitted with a Cohn[30, 31] salting-out equation analogous, and selected experiments were compared with the theoretical models from Sim et al. ,[32] Odijk[33] and Atha and Ingham.[18]

2. MATERIALS AND METHODS

2.1. GLYCOSYLATED MIXTURE AND NON-GLYCOSYLATED MAB

The mAbs were generously supplied by Monsanto Protein Technologies (St. Louis, MO) or extracted from recombinant corn flour according to the procedure described in Lee and Forciniti[34] and purified using a Protein A column. GSL-II, and PEG of 8000 and 1450 Da were purchased from Sigma (St. Louis, MO). All the other chemicals were of analytical grade. Two protein preparations were used: 1) a fully non-

glycosylated mAb (we refer to this preparation as non-glycosylated) and 2) a mixture of 25% non-glycosylated, 50% single GlcNAc glycosylated mAb, and 25 % fully glycosylated mAb (we refer to this preparation as glycosylated mixture).

2.2. PRECIPITATION EXPERIMENTS

0.31g of the glycosylated mixture or 0.61g of the non-glycosylated protein mAb (or a total of 1.27mg each protein) were mixed in a centrifuge tube with various amounts of a PEG stock solution (50% w/w of either PEG 1450 or PEG 8000) and phosphate buffer at pH 6, TRIS buffer at pH 9, acetate buffer at pH 4, or carbonate-bicarbonate buffer at pH 10.7 to complete 2.5 g. GSL II was added to some systems at a concentration (in moles) similar to that of the mAb. The contents of the tubes were mixed in an orbital mixer for 20 minutes and then centrifuged at either 4 °C, 25 °C, or 40 °C for 90 minutes at 2,800 ×g. Samples of the supernatant were removed and absorbance at 280 nm was measured in a double beam spectrophotometer. Appropriate blanks were prepared each time and their absorbance values were discounted from the samples' reading. For the systems containing lectins, the precipitate was re-suspended in the appropriate buffer and run through a gel permeation column using a UV detector in tandem with the MALS detector.

2.3. MULTIPLE ANGLE LASER LIGHT SCATTERING (MALS)

Samples of the glycosylated and non-glycosylated mAb were injected through a 50 μL or a 250 μL loop into a gel permeation column (Protein KW-803, Shodex). Two detectors were connected in series: 1) a Hitachi UV spectrophotometer and 2) a Wyatt

multiple angle laser light scattering detector. The flow rate was either 1 ml/min or 0.5 ml/min. The same procedure was followed with re-suspended precipitates of the antibodies in the presence of lectins.

The extinction coefficients of the glycosylated and non-glycosylated mAbs were determined as ε= 1.37 cm^2/mg and ε = 1.15 cm^2/mg, respectively.[34] The molecular weight of the proteins was determined by the Debye plotting of the scattering data using an internal calibration constant for the UV detector.

2.4. EXPERIMENTAL DESIGN

A full two-level experimental design was used. The factors included pH (6 and 9), glycosylation of the mAb (glycosylated and non-glycosylated) and PEG molecular weight (8,000 and 1450). Additional experiments were performed at 5, 25 and 40 ^oC and in the presence of GSL-II.

2.5. THEORETICAL BACKGROUND

A semi-logarithmic analogous to the Cohn salting-out equation was suggested by Juckes[31] to correlate protein solubility with PEG concentration,

$$\log S = \beta\omega + \kappa \qquad (1)$$

where S is protein solubility, ω is the PEG concentration in % w/w, and κ is the log of the intrinsic solubility of the protein in the absence of PEG. The β-value that represents the precipitation efficiency can be obtained by fitting the linear region of the solubility data with Eq.1.

Differentiating Eq. 1 with respect to the polymer concentration yields,

$$\frac{d\log S}{d\omega} = \frac{\Delta \log S}{\Delta \omega} = \beta \tag{2}$$

Using the depletion theory of nanoparticles immersed in a semi dilute polymer solution, Odijk[33] derived an expression for β in terms of hydrodynamic radius of the protein ($r_{h,prot}$),

$$\frac{\Delta \log S}{\Delta \omega} = \beta = -0.036 r_{h,prot} \tag{3}$$

Another model for β was proposed by Atha and Ingham.[18] Their model is based on excluded volume considerations and yields the following expression for β:

$$-\beta = \frac{a}{0.23 M_2} \tag{4}$$

where M_2 is the molecular weight of PEG, and a is the interaction coefficient between protein and polymer, which can be calculated by the expression of molar excluded covolume for pairs of spherical molecules 2-3 as proposed by Ogston,[35]

$$U_{23} = 10^3 a = \frac{4\pi N}{3}(r_{h,prot} + r_{h,PEG})^3 \tag{5}$$

where U_{23} is the molar excluded covolume for pairs of spherical molecules 2-3, N is Avogadro's number, and $r_{h,PEG}$ is the hydrodynamic radius of PEG.

Based on previous works,[18, 33, 36-38] Sim *et al.* [32] proposed a model for β that includes the hydrodynamic radius of protein ($r_{h,prot}$) and PEG ($r_{h,PEG}$),

$$\beta = (\gamma \cdot r_{h,PEG}^{0.211} + \delta) r_{h,prot} \tag{6}$$

where γ and δ are regression parameters. The first term ($\gamma \cdot r_{h,PEG}^{0.211} r_{h,prot}$) was interpreted by Sim *et al.* as the depletion of protein by PEG whereas the second term (δ $r_{h,prot}$) accounts

for the volume excluded by the protein and the depletion of PEG by the protein. This depletion of PEG by protein allows the protein to remain soluble. They regressed data for 6 different proteins with PEGs of three molecular weights at each protein's isoelectric point.[39] Regression of the experimental data sets yielded the following correlation for β with the radius of the polymer and the protein,[39]

$$\beta = (0.076 r_{h,PEG}^{0.211} - 0.045) r_{h,prot} . \quad (7)$$

The regression was not very good as evidenced by a correlation coefficient of R^2=0.725. Compared to Atha and Ingham's model, Sim's model is a simple correlation for protein solubility with PEG and protein sizes, but it is based on fundamental principles of polymer physics.

All three models need the hydrodynamic radius of protein ($r_{h,prot}$) and PEG ($r_{h,PEG}$) to calculate the β-value. In this study the $r_{h,PEG}$ was calculated using the correlation by Fee *et al.*,[36]

$$r_{h,PEG} = 0.1912 M_{r,PEG}^{0.559} \quad (8)$$

The $r_{h,prot}$ of both proteins was determined using the molecular weights obtained from the *MALS* results assuming globular proteins (molecular weight: 150 kDa and $r_{h,prot}$= 5.3 nm).

3. RESULTS AND DISCUSSION

We first determined the precipitation of the mAb solutions in the absence of PEG. Measurements of protein concentration after centrifugation of the standard solutions were done after 20 min., 2.5, 12, 24, and 36 hours for selected systems (pH 4, pH 6, pH 10.7,

and pH 6 with GSL II) in the absence of PEG. The decrease in the solubility of the glycosylated sample was between 0.1 to 7% after 36 hrs. On the contrary, the decrease in the solubility of the non-glycosylated mAb ranged from 0.1% (pH 6 plus GSL II) to 40 % (pH 10) also after 36 hrs. The presence of the lectin seems to stabilize the non-glycosylated mAb. The isoelectric point of the non-glycosylated mAb is from 9.7 to 9.1 whereas the isoelectric point of the glycosylated mixture expands the range from 9.4 to 8.7[34, 40]. Near its isoelectric point, the non-glycosylated mAb solution becomes quite unstable after 24 hours. The experiments containing PEG were done using an incubation time of 20 minutes.

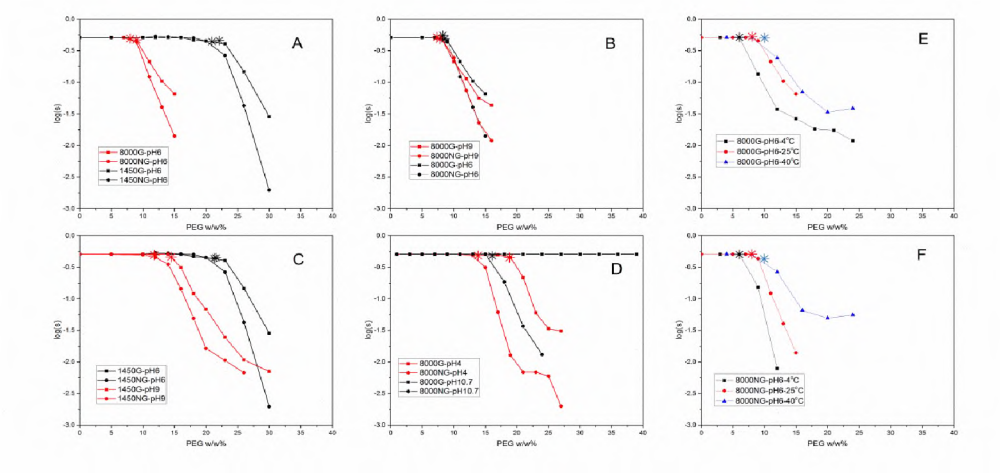

Figure 1. Effect of PEG molecular weight, pH and temperature on the solubility of mAbs. 'G' or 'NG' represent glycosylated mixture and non-glycosylated mAbs in the legend and marks ✳ is the onset precipitation concentration of PEG. Panel A. effect of PEG molecular weight. Panel B, pH effect in PEG 8000. Panel C, effect of pH at PEG 1450. Panel D, effect of pH at pHs which are more away from pIs. Panel E and F, effect of temperature

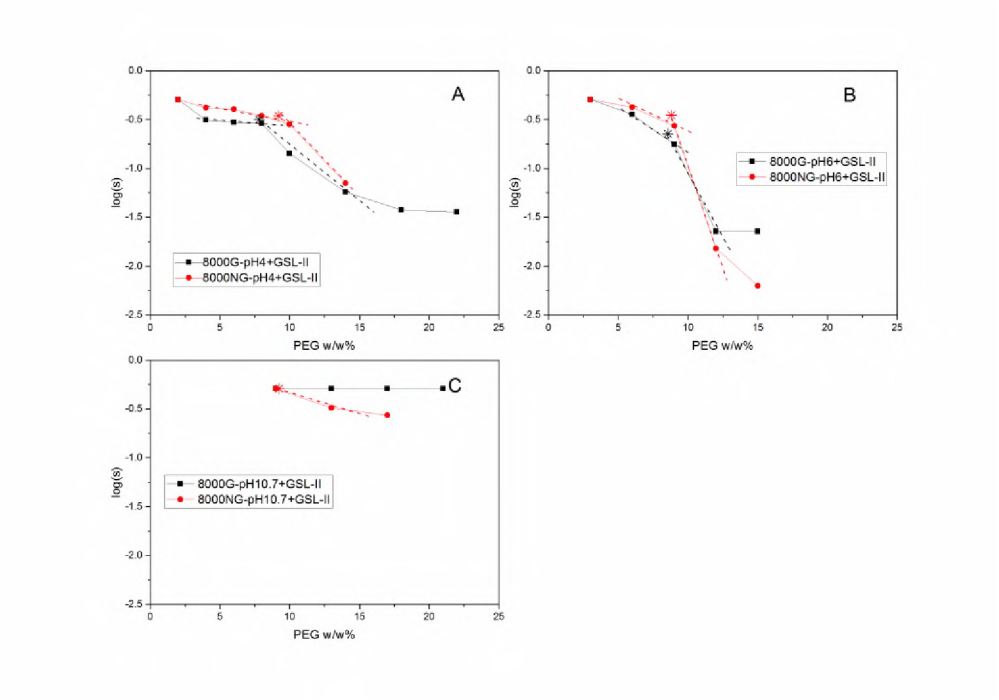

Figure 2. Effect of Griffonia (Bandeiraea) simplicifolia Lectin II (GSL II) on the solubility of mAbs upon addition of PEG 8000. The dashed lines are used to show a two-step precipitation process. Panel A: pH 4. Panel B: pH 6. Panel C: pH 10.7

Plots of antibody solubility vs. polymer concentration at different conditions are shown in Figures 1 and 2. “G” and “NG” in the figures represent the glycosylated mixture and the non-glycosylated mAb, respectively. Figure 1 shows the effect of pH, PEG molecular weight and temperature and Figure 2 shows the effect of the addition of a lectin on the solubility of the mAbs. Figure 1A shows that both preparations are more soluble in PEG 1450 than PEG 8000. Both preparations are more soluble at pH 6 than 9 in the presence of either PEG 1450 or PEG 8000. At very acidic (pH 4) or very basic (pH 10.7) pHs the solubility of both preparations increases but the solubility of the glycosylated mAb is more sensitive to pH changes. The results obtained at pH 10.7 are

striking since the glycosylated mAb does not precipitate in the PEG concentration range covered by these experiments, in spite of the fact of containing 25% of the non-glycosylated protein. At the highest PEG concentration used to precipitate the non-glycosylated mAb (24% w/w) approximately 25% of the antibody remains in solution; which may explain the absence of precipitation at pH 10.7. Figure 1E and F show the effect of temperature on the solubility of the mAb in the presence of PEG. At 4°C and at a PEG concentration of 12% only 1% of the non-glycosylated mAb remains in solution whereas about 8% of the glycosylated mixture remains in solution at the same conditions. Finally, Figure 3 shows that the precipitation of the mAbs in the presence of a lectin follows a two-step process at pH 4 and 6. The precipitation curves were fitted with a Cohn salting-out equation analogous (Eq.1) to obtain the precipitation efficiency coefficient, β. The m*-values (the PEG concentration at the onset of precipitation) where calculated from the intercept of the Cohn equation with a horizontal line at the initial protein concentration.

3.1. m* VALUES

The m* values for all runs are shown in Table 1 and in Figure 3. Figure 3A shows that, with a few exceptions, the value of m* is larger for the glycosylated than for the non-glycosylated mAb. The difference in the m*-value of the glycosylated mixture and non-glycosylated mAb is more pronounced at pH 4 and PEG 8000 (Figure 3B and C). This confirms that one of the roles played by the carbohydrates chains is to increase protein solubility. The m*-values of both preparations decrease as the temperature decreases as expected.

Table 1. m*-values and β-values of glycosylated mAb mixture and non-glycosylated MAb precipitation curves

PEG/Da	pH	T/°C	Additive	Glycosylated mixture		Non-Glycosylated MAb	
				β-value	m*-value	β-value	m*-value
1450	6	25	N/A	-0.16±0.008	22.46±0.2	-0.31±0.02	21.6±0.2
1450	9	25	N/A	-0.14±0.007	14.1±0.1	-0.22±0.008	13.31±0.1
8000	4	25	N/A	-0.20 ± 0.02	18.8±0.8	-0.283 ± 0.004	13.9±0.1
8000	6	25	N/A	-0.14 ± 0.01	8.4±0.2	-0.248 ± 0.007	8.7±0.3
8000	9	25	N/A	-0.12 ± 0.02	8.0±0.3	-0.22 ± 0.02	8.4±0.4
8000	10.7	25	N/A	0	N/A	-0.19 ± 0.02	15.3±0.9
8000	6	4	N/A	-0.184	6.09	-0.429	7.78
8000	6	40	N/A	-0.11 ± 0.02	8.6±0.4	-0.153	9.96
				1st 2nd	1st 2nd	1st 2nd	1st 2nd
8000	4	25	GS L II	-0.0086 ± 0.002	-0.087 ± 0.01	2	7.9 ±0.4	-0.021 ± 0.001	-0.12 ± 0.03	2.9 ±0.09	8.2 ±0.5
8000	6	25	GS L II	-0.1	-0.2	4.5	7.1 1	-0.064	-0.24 1	5.1	6.8 4
8000	10.7	25	GS L II	0	0	N/A	-0.034 ± 0.009	N/A	9.0 ±0.2	N/A

To identify the main effects on the m*-values, a Paretto chart (Figure 3D) was constructed using three experimental factors: 1) type of mAb, 2) PEG molecular weight and 3) pH. The most significant factor is the molecular weight of PEG whereas the presence of the glycan groups is less statistically significant on the onset of precipitation. Recently, a quantitative structure-activity relationship (QSAR) [41] model was used to correlate the experimental m*-values and β-values with 132 molecular descriptors. The coefficient of determination R^2 for m*-values is 0.9 and for β-values is 0.93.

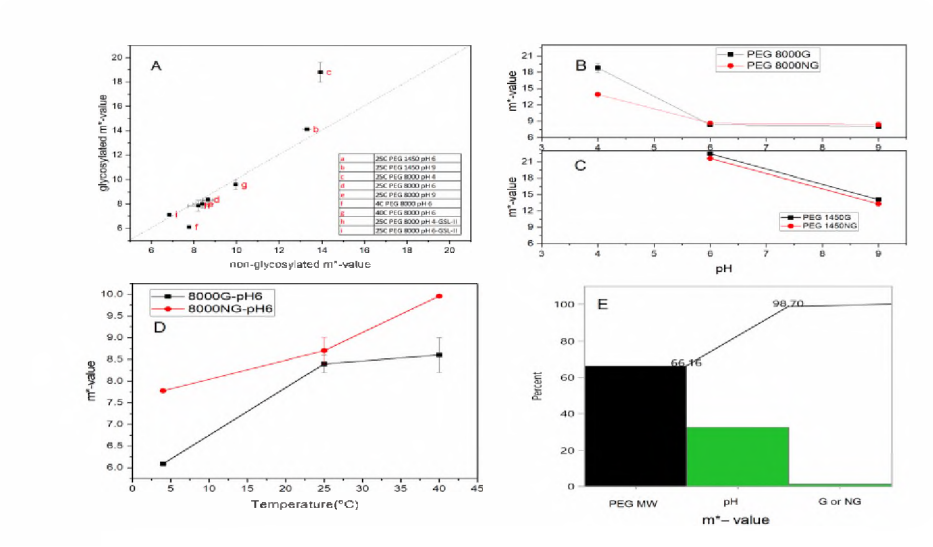

Figure 3. m*-values.Panel A. Comparison of m*-values of both mAb preparations. Panels B (PEG Mw 8000) and C (PEG Mw1450). pH dependence of the m*-values; ■, glycosylated mAb; ●, non-glycosylated mAb. Panel D. Paretto chart for m*-values. PEG molecular weights: 1450 or 8000Da, pHs: 4 or 6, and both types of mAbs. Panel E, Effect of temperature on m*-value

That study paves the way to explain protein precipitation by PEG based on protein surface properties. The 10 molecular descriptors with highest variable of the influence on the projection (VIP) of m*-values are shown in Table 2 in decreasing order of importance. The direction of the influence of each factor on m*-value is given by the sign of the regression coefficient i.e., an increase in a molecular descriptor with a positive (+) regression coefficient will increase the amount of PEG per protein needed to start protein precipitation and vice versa. [41] They concluded that there are four major factors that can influence m*-values: the sphericity of the protein (-)> density of the protein (+)> electrostatic surface potential (ESP) (+)> solvent accessible surface area of protein (+). The ESP is easily manipulated by changing the pH of the solution. Still, there may be an

indirect effect of the carbohydrate chain on the ESP since its presence may affect the dissociation constant of ionizable amino acids side chains. The ESP of a protein will increase as the solution pH is farther away from the protein's isoelectric point. Therefore, the amount of PEG needed to precipitate the protein will increase as the ESP increases, which agrees with our finding for both preparations with PEG 8000 and PEG 1450. The effect of pH for the whole design is somehow masked by the dominance of PEG molecular weight. The sphericity and density of the proteins barely change upon glycosylation. However, glycosylation does increase the solvent accessible surface area of protein.[42] Because the regression coefficient for solvent accessible surface area is positive and the glycosylated protein has a larger surface area, the onset of precipitation should happen at a higher PEG concentration for the glycosylated sample, as observed. However, the effect is not statistically significant at 95% confidence.

Table 2. Descriptors with a VIP value >1.0 included in the final QSAR model for m*-value and their description[a]

No.	Descriptor	Definition	Sign of Regression coefficient
1	shapeMin	Value for the sphericity of the protein:(minimum distance between mass center and protein surface)/(mean distance between mass center and protein surface)	-
2	dens	Density of the protein	+
3	sumSurfA_ShellEsp	Sum of ESP of surface points projected on a shell around the molecule with a distance of 5Å	+

Table 2. Descriptors with a VIP value >1.0 included in the final QSAR model for m*-value and their description[a] (cont.)

No.	Descriptor	Definition	Sign of Regression coefficient
4	totalSurf_patchEsp	Solvent-accessible surface area of protein in $Å^2$ on the patch with the highest ESP value	+
5	toalSurfA_Shell	Solvent-accessible surface area of a shell around the molecule with a distance of 5 Å	+
6	totalSurf_PatchHyd	Solvent-accessible surface area of the protein surface patch with the highest hydrophobicity value in $Å^2$	+
7	nAAcid	Chain length of the protein	+
8	mAtom	Number of atoms of the protein	+
9	mass	Molecular mass of the molecule	+
10	devA_PlaneEsp	(maximum ESP value-minimum ESP value)/mean value of ESP on the plane with the highest ESP value	-

[a]Reproduced with permission from reference 33

3.2. β-VALUES

The *β*-values and associated errors of all runs are presented in Table 1 and plotted in Figure 4. The number of data points in some of the runs is too small and therefore errors in the fitting cannot be calculated. In all runs (even in the lectin containing runs) the *β*-values are more negative for the non-glycosylated mAb than for the glycosylated mAb, which indicates that PEG is more efficient precipitating the non-glycosylated mAb. At the same pH, the higher the molecular weight of PEG the higher the efficiency to precipitate both the glycosylated and non-glycosylated mAbs. At the same PEG molecular weight, the efficiency decreases with increasing pH (Figure 4B and C). The *β*-

values for the glycosylated mAb at pH 10.7 with the addition of PEG 8000 was set to zero because the glycosylated mAb cannot be precipitated in the PEG concentration studied. In the presence of PEG 8000 at pH 6, the efficiency increases with decreasing temperature for both samples (Figure 4E) but the solubility of the non-glycosylated mAb is more sensitive to changes in temperature than the glycosylated one(Figure 5E). For example, at 12% PEG, nearly 70% of the non-glycosylated protein remains in solution at 40 °C whereas only 1% remains in solution at 4 °C.

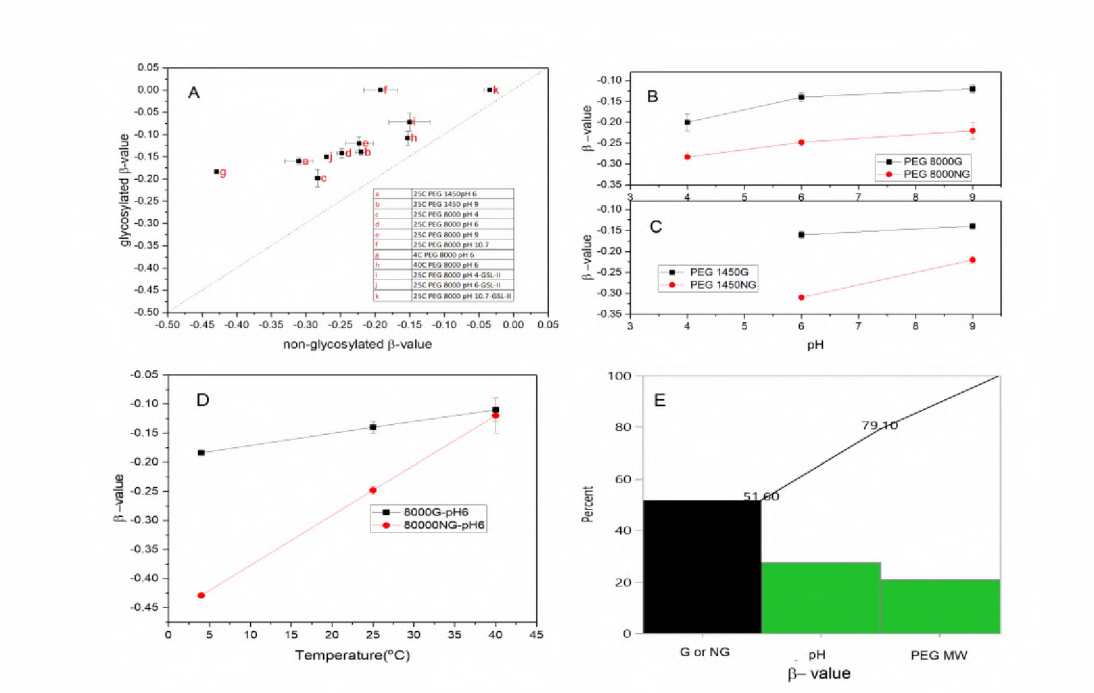

Figure 4. β-values for the entire experimental base. Panel A, comparison of β-values of the glycosylated and non-glycosylated mAb. ■, glycosylated mAb; ●, non-glycosylated mAb. Panels B (PEG 8000) and C (PEG 1450). β-values of glycosylated and non-glycosylated mAb vs. pH; ■, glycosylated mAb; ●, non-glycosylated mAb. Panel D: Paretto chart of β-values. PEG molecular weights: 1450 or 8000Da; pHs: 4 or 6. Panel E: effect of temperature on β-values. ■, glycosylated mAb; ●, non-glycosylated mAb

A Paretto chart was constructed using three experimental factors: 1) type of mAb, 2) PEG molecular weight and 3) pH. Figure 4D shows that the effect of those factors on β is in the order: glycosylated or non-glycosylated> pH >PEG molecular weight. This order is opposite to the order observed for the onset of precipitation. The Paretto chart shows that the effect of the presence of carbohydrates chain on β is more important than the effects of pH (associated with electrostatic repulsions). Therefore, although glycosylation has a moderate effect on the onset of precipitation it has a dominant effect on the precipitation efficiency by PEG.

The trends in the β-values of all runs without GSL-II were analyzed in the context of the three models introduced earlier (Figure 5). The hydrodynamic radii of PEG used in the models were calculated using Eq.8 whereas the hydrodynamic radii of the proteins were measured using MALS. All three models show that the β-values increase with an increase in the protein hydrodynamic radius and PEG molecular weight.

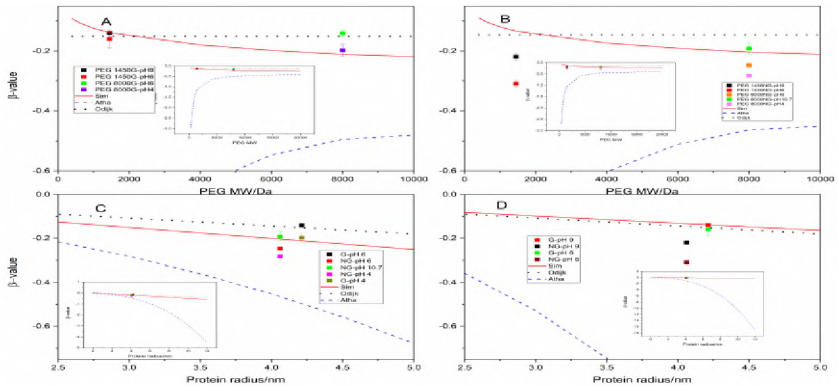

Figure 5. Comparison of β-values of selected runs with three predictive models.Panels A (glycosylated mAb) and B (non-glycosylated mAb): β-values vs. PEG MW. Panels C (PEG MW 1450 Da) and D (PEG MW 8000 Da): β-values vs. protein radius. Inserts are zoom outs of the plots

The predicted β-values of the glycosylated and non-glycosylated mAbs in Odijk's and Sim's models are similar to the experimental ones, but Atha and Ingham's model overestimates those β-values. The glycosylated mAb has a slightly larger molecular weight/hydrodynamic radius, which should yield higher β-values. Both Odijk's and Atha's models are based on excluded volume considerations. Whereas Atha's model assumes that protein and polymer are hard bodies (spheres and cylindrical fibers) Odijk's model assumes that the polymer is a chain of monomers and that particles are of arbitrary shape but have an equivalent radius of gyration. By incorporating the hydrodynamic radius of PEG and the Stokes radius of the protein, Sim's model provides a quick guideline to select precipitation conditions by PEG. Our experimental data show that pH has an effect on the β-values that cannot be explained by any of the three models. Furthermore, all three models predict an opposite trend for the β-values of glycosylated and non-glycosylated mAbs. Although the carbohydrate chain on the glycosylated mAb surface and its associated water shell increase the hydrodynamic radius, it does not promotes precipitation. On the contrary, the carbohydrate chain increases protein's solubility. We argue that these models fail because they do not account for electrostatic repulsions/attraction and for hydration forces (enhanced by the presence of glycan groups). Enhanced hydration may justify the higher stability of the glycosylated mAb.

We also compared the β-values with the predictions of the QSAR model. The 19 molecular descriptors with highest variable influence on the projection (VIP) values are shown in Table 3 in descending order. As before, the direction of the influence of each factor on β-values is given by the sign of the regression coefficient i.e., an increase in a molecular descriptor with a negative (-) regression coefficient makes the precipitation

curve steeper and vice versa. Hammerling *et al.* [41] found that the three most effective factors affecting β are solvent accessible surface area of protein (-) > protein molecular weight (-), number of atoms and chain length of the protein (-) > electrostatic surface potential (-).

Table 3. Descriptors with a VIP value >1.0 included in the final QSAR model for β-value and their description[a]

No.	Descriptor	Definition	Sign of Regression coefficient
1	totalSurf	Solvent accessible surface area of protein in Å^2	-
2	mass	Molecular weight of the molecule	-
3	nAtom	number of atoms of the protein	-
4	nAAcid	Chain length of the protein	-
5	shapeFactor	Value for the sphericity of the protein	-
6	sumNeg_PatchEsp	Sum of negative ESP on the protein patch with the highest ESP value	+
7	sumNeg_SurfEsp	Sum of negative ESP on the protein surface	+
8	median_PlaneESP	Median value of ESP on the protein patch with the highest ESP value	-
9	sumSurf_PatchEsp	sum of ESP of surface points on the protein patch with the highest ESP value	+
10	sum_SurfEsp	Sum of ESP of surface point on the protein surface	+
11	totalSurf_PatchEsp	Solvent accessible surface area of the protein patch with the highest ESP value	-
12	ninAbs_SurfHyd_4	Number of points with low hydrophobicity on the protein surface	-
13	sumPos_SurfHyd	Sum of points with positive hydropathy score on the protein surface	-

Table 3. Descriptors with a VIP value >1.0 included in the final QSAR model for β-value and their description (cont.)

No.	Descriptor	Definition	Sign of Regression coeffic ient
14	devA_PatchHyd	(maximum hydrophobicity value-minimum hydrophobicity value)/mean value of hydrophobicity on the patch with the highest hydrophobicity	-
15	binAbs_SurfHyd_3	Number of points with low hydrophobicity on the protein surface	-
16	charge	overall charge of protein	+
17	median_ShellEsp	Median value of ESP projected on a shell around the molecule with a distance of 5 Å	-
18	max_PatchHyd	Maximum value hydrophobicity on the protein patch with the highest hydrophobicity value	+
19	sumNeg_SurfHyd	Sum of negative hydrophobicity values on the protein surface	+

[a]Reproduced with permission from reference 33

These trends agree with Atha and Ingham's[18] findings that larger proteins will have a steeper slope; i.e., the larger the molecular weight of the protein the higher the possibility of protein-protein interaction. This conclusion is contradicted by the findings of Sola *et al.* [43] who found that glycosylation enhances the stability of α-chymotrypsin because glycosylation increases the solvent accessible surface area linearly. In the QSAR model, the increased solvent accessible surface area is the result of an increasing number of hydrophilic amino acids in the polypeptide chain whereas in Sola *et al.*'s experiments

it is caused by glycosylation. Glycosylation will increase the solvent accessible surface area without affecting the size of the protein as much as hydrophilic amino acids. For example, arginine, the most hydrophilic amino acid residue, increases the solvent accessible surface area by ~1.2 $Å^2$/Da[44] but a typical glycan group like lactose will increase the area by ~1.5 $Å^2$/Da. The QSAR model cannot capture the carbohydrate chain's effects on PEG-induced precipitation. This should not be surprising since the model does not include the same protein with and without glycan groups in the study and the glycan groups of the glycoproteins were not included in the MD simulation. The glycosylated mAb used in our experiments has a larger molecular weight and a larger solvent accessible surface area than the non-glycosylated mAb, but it has a flatter slope than the non-glycosylated counterpart, which contradicts the QSAR model. We argue that the solvent accessible surface area itself should not be used as one of the indicators of protein stability. Instead, solvent mediated forces, like hydration ones, would be a better predictor. The presence of glycan groups in a protein stabilizes the polypeptide chain by increasing its hydration. Protein stabilization caused by hydration forces is, in part, compensated by the destabilization caused by an increase in protein molecular size.

3.3. ADDITION OF LECTINS

Figure 2 shows that the addition of GSL-II produces a two-stage precipitation pattern at pH 4 and 6. The pattern consists of a first stage in which there is gentle decrease in the solubility followed by a second stage in which the solubility decreases sharply (Figure 2 dash line). Two mechanisms may be considered here: 1) GSL-II is not specific to non-glycosylated mAb, but it still lowers the m*-value sharply, which implies

that the non-glycosylated mAb co-precipitated with GSL-II by adding PEG. The enhanced co-precipitation by PEG or hetero-interaction among variant proteins was reported by Miekka and Ingham.[45] They argue that it only occurs when the hetero-complex is already present in the solution before PEG is added (proteins forming hetero-complex before adding PEG). 2) GSL-II specifically binds to the mono-glycosylated mAb in the mixture, and that is the species that precipitates. Co-precipitation may also occur in the glycosylated mixture since 25% of mAb is non-glycosylated. Miekka and Ingham[45] also pointed out that the electrostatic interactions between proteins are the main reason of forming hetero-complex; i.e., the enhancement of co-precipitation in binary mixtures was a maximum at the pH intermediate between the two isoelectric points of the proteins (the proteins carried opposite net charge). Considering the isoelectric point of the mAb and the lectin, the hetero-complex should be more stable at pH 6 than at pH 4. In the next two paragraphs, the second precipitation stage is discussed in more detail.

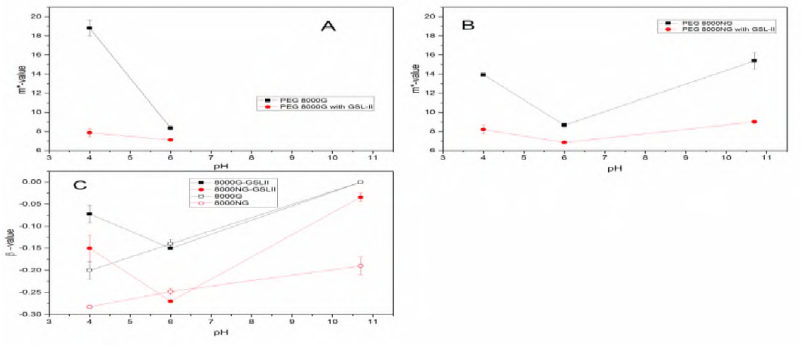

Figure 6. The m*-values and β-values vs. pH for the precipitation curves with GSL-II and PEG 8000.Panels A (glycosylated mAb) and B (non-glycosylated mAb): comparison of the m*-values of precipitation curves with or without GSL-II. Panel C: comparison of the β-values of glycosylated mAb and non-glycosylated mAb in the presence and absence of GSL-II

The addition of GLS-II decreases the m*-value of both preparations (Figure 6A and B) and the effect is more pronounced at pH 6 where both the mAb and the lectin (pI: 5.6~6) are positively charged. More importantly, pH 6 is in the middle the optimum binding pH range of GSL-II and GlcNAc is 4.2 ~8.8.[46] At the same pH (pH 4 and pH 6), the m*-values of the glycosylated mAb are slightly smaller than the ones for the non-glycosylated mAb's in the presence of GSL-II. Still, at pH 10.7, the glycosylated mixture cannot be precipitated even in the presence of GLS-II. pH 10.7 is outside GSL-II's optimum binding pH range; and therefore its presence did not affect the anomalous behavior observed with the pure mAb at this pH.

The first stage absolute β-values are smaller than the second stage β-values (Table 1). The second stage β-values are around -0.087~ -0.12 for pH 4 and -0.20~ -0.24 for pH 6 with and without GSL-II, respectively. The second stage β-values for both mAbs in the presence of GSL-II show a minimum at pH 6, which indicate that the PEG precipitation efficiency is highest at that pH. The enhancement of precipitation efficiency at pH 6 for both mAbs may be caused by the proximity of that pH to the isoelectric point of GSL-II (Ip 5.6~6.0). Because GSL-II also binds to GlcNAc, the differences in β-values of the glycosylated mAb with and without GSL-II are larger than for the non-glycosylated one. GSL-II lowers the PEG precipitation efficiency of the non-glycosylated mAb at all pHs but especially at pHs 4 and 10.7. Both the lectin and the mAb are either positively (pH 4) or negatively charged (pH 10.7) at those pHs and therefore electrostatic repulsions are strengthened. The same observation can be made for the glycosylated mAb at pH 4 and 10.7. Once again, even in the presence of the lectin the glycosylated mAb does not precipitate in the range of PEG concentrations used in this study. GSL-II decreases the

differences in the precipitation efficiency in between both preparations at all pHs. This is expected because the stability provided by the carbohydrate chains is lost as the lectin binds the mAb.

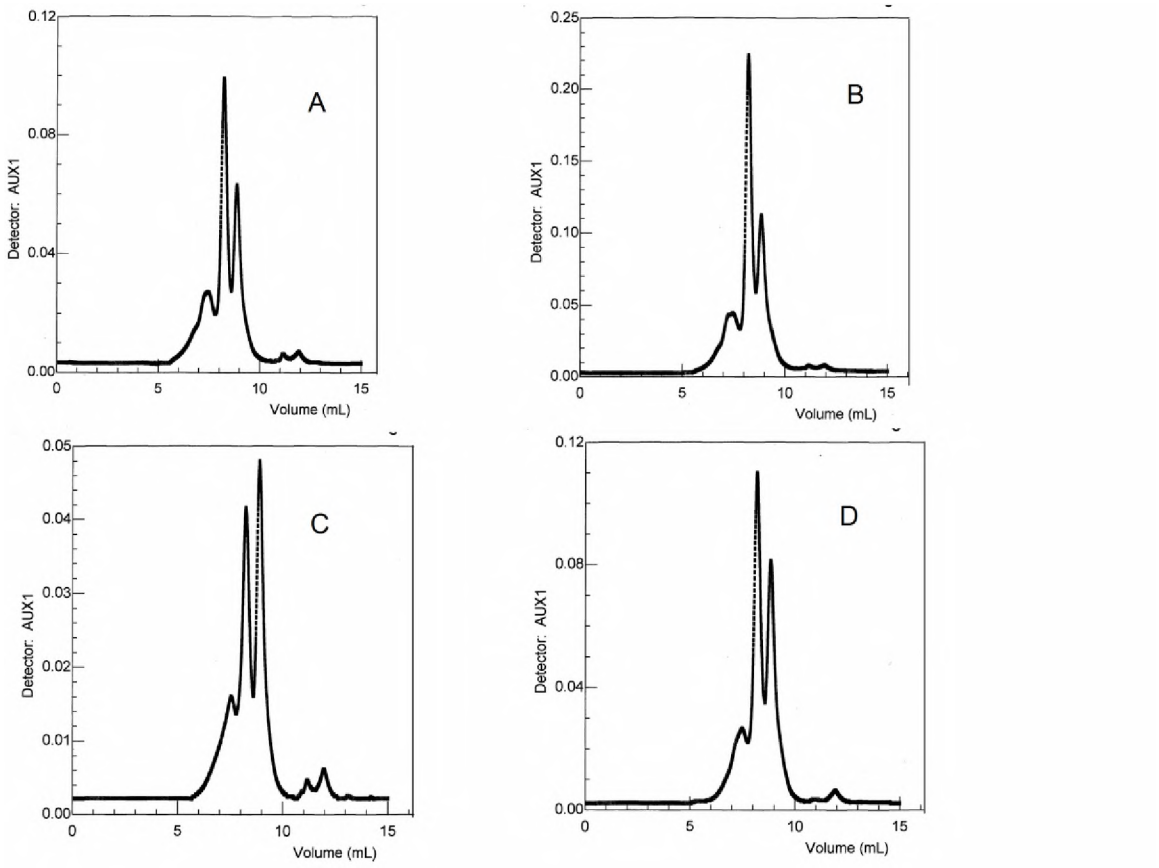

Figure 7. Chromatograms of the re-suspended precipitates in the presence of a lectin.Panel A: non-glycosylated mAb at 16.6 % precipitation. Panel B: glycosylated mAb at 29.6 % precipitation. Panel C: non-glycosylated mAb at 46.3 % precipitation. Panel D: glycosylated mAb at 65.3% precipitation

There are some differences in the precipitation pattern of the non-glycosylated protein and the glycosylated mixture and between the composition of the precipitate at the beginning of the precipitation and at the end (Figure 7) in the presence of GSL-II. At

the onset of precipitation (Figures 7 A and B) there are no major differences in the composition of the non-glycosylated and glycosylated mAb precipitates. However, in the middle of the precipitation curve (Figures 7 C and D) the precipitate of the non-glycosylated mAb is richer in the lectin (peak at ~ 9 ml) than the corresponding precipitate of the glycosylated one. Base on the HPLC-MALS results and recalling the co-precipitation mechanism, the first stage may represent the precipitation of hetero-complex which is larger than glycosylated or non-glycosylated mAb, the second stage represents the free mAbs or the GSL-II-GlcNAc-mAb complex. For both systems C and D there is a portion of the precipitate that cannot be re-suspended. We speculate that this precipitate consists of aggregates of the lectin with the GlcNAc glycosylated mAb. The two-stage pattern disappeared in the non-glycosylated mAb precipitation curve at pH 10.7, which implies the hetero-complex cannot be formed because both of the non-glycosylated mAb and GSL-II are negatively charged.

4. CONCLUSIONS

The results presented in this paper demonstrate the differences in the solubility of the glycosylated and non-glycosylated mAbs. PEG has lower precipitation efficiency and a larger onset concentration for the non-glycosylated than for the glycosylated mixture. This manuscript demonstrates that depletion and electrostatic forces are not sufficient to explain protein precipitation by PEG. The precipitation experiments with lectin show that there is a complex interplay between the mAbs and GSL-II. This is demonstrated by the analysis of the precipitates in the presence of a lectin. Our studies also show that

available models to correlate or predict protein solubilities are unable to capture the effect of glycosylation on protein solubility.

REFERENCES

[1] Ecker, D. M., Jones, S. D., Levine, H. L., The therapeutic monoclonal antibody market. *MAbs* 2015, *7*, 9-14.

[2] Maverakis, E., Kim, K., Shimoda, M., Gershwin, M. E., *et al.*, Glycans in the immune system and The Altered Glycan Theory of Autoimmunity: A critical review. *J. Autoimmun.* 2015, *57*, 1-13.

[3] Jennewein, M. F., Alter, G., The Immunoregulatory Roles of Antibody Glycosylation. *Trends Immunol.* 2017, *38*, 358-372.

[4] Spencer, S., Bethea, D., Raju, T. S., Giles-Komar, J., Feng, Y., Solubility evaluation of murine hybridoma antibodies. *MAbs* 2012, *4*, 319-325.

[5] Pepinsky, R. B., Silvian, L., Berkowitz, S. A., Farrington, G., *et al.*, Improving the solubility of anti-LINGO-1 monoclonal antibody Li33 by isotype switching and targeted mutagenesis. *Protein Sci.* 2010, *19*, 954-966.

[6] Mitra, N., Sinha, S., Ramya, T. N. C., Surolia, A., N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem. Sci.* 2006, *31*, 156-163.

[7] Rademacher, T. W., Parekh, R. B., Dwek, R. A., GLYCOBIOLOGY. *Annu. Rev. Biochem.* 1988, *57*, 785-838.

[8] Karpusas, M., Whitty, A., Runkel, L., Hochman, P., The structure of human interferon-β: implications for activity. *Cell. Mol. Life Sci.* 1998, *54*, 1203-1216.

[9] Masarova, J., Mislovicova, D., Gemeiner, P., Michalkova, E., Stability enhancement of Escherichia coli penicillin G acylase by glycosylation with yeast mannan. *Biotechnol. Appl. Biochem.* 2001, *34*, 127-133.

[10] Nissen, C., Glycosylation of recombinant human granulocyte colony stimulating factor: implications for stability and potency. *Eur. J. Cancer* 1994, *30A Suppl 3*, S12-14.

[11] Narhi, L. O., Arakawa, T., Aoki, K. H., Elmore, R., *et al.*, The effect of carbohydrate on the structure and stability of erythropoietin. *J. Biol. Chem.* 1991, *266*, 23022-23026.

[12] Barbarić, S., Mrša, V., Ries, B., Mildner, P., Role of the carbohydrate part of yeast acid phosphatase. *Arch. Biochem. Biophys.* 1984, *234*, 567-575.

[13] Oh-eda, M., Hasegawa, M., Hattori, K., Kuboniwa, H., *et al.*, O-linked sugar chain of human granulocyte colony-stimulating factor protects it against polymerization and denaturation allowing it to retain its biological activity. *J. Biol. Chem.* 1990, *265*, 11432-11435.

[14] Harris, R. J., Shire, S. J., Winter, C., Commercial manufacturing scale formulation and analytical characterization of therapeutic recombinant antibodies. *Drug Dev. Res.* 2004, *61*, 137-154.

[15] Wang, X., Das, T. K., Singh, S. K., Kumar, S., Potential aggregation prone regions in biotherapeutics. *MAbs* 2009, *1*, 254-267.

[16] Moussa, E. M., Panchal, J. P., Moorthy, B. S., Blum, J. S., *et al.*, Immunogenicity of Therapeutic Protein Aggregates. *J. Pharm. Sci.* 2016, *105*, 417-430.

[17] Wang, W., Instability, stabilization, and formulation of liquid protein pharmaceuticals. *Int. J. Pharm.* 1999, *185*, 129-188.

[18] Atha, D. H., Ingham, K. C., Mechanism of precipitation of proteins by polyethylene glycols. Analysis in terms of excluded volume. *J. Biol. Chem.* 1981, *256*, 12108-12117.

[19] Iverius, P. H., Laurent, T. C., Precipitation of some plasma proteins by the addition of dextran or polyethylene glycol. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* 1967, *133*, 371-373.

[20] Mitchison, T. J., Colloid osmotic parameterization and measurement of subcellular crowding. *Mol. Biol. Cell* 2019, *30*, 173-180.

[21] Thrash, S. L., Otto, J. C., Deits, T. L., Effect of divalent ions on protein precipitation with polyethylene glycol: Mechanism of action and applications. *Protein Expr. Purif.* 1991, *2*, 83-89.

[22] Dao-Thi, M.-H., Hamelryck, T. W., Poortmans, F., Voelker, T. A., *et al.*, Crystallization of glycosylated and nonglycosylated phytohemagglutinin-L. *Proteins* 1996, *24*, 134-137.

[23] Pauly, D., Chacana, P. A., Calzado, E. G., Brembs, B., Schade, R., IgY technology: extraction of chicken antibodies from egg yolk by polyethylene glycol (PEG) precipitation. *J. Vis. Exp.* 2011, 3084-3090.

[24] Boncina, M., Rescic, J., Vlachy, V., Solubility of lysozyme in polyethylene glycol-electrolyte mixtures: the depletion interaction and ion-specific effects. *Biophys. J.* 2008, *95*, 1285-1294.

[25] Zhao, Y., Kang, L., Gao, S., Gao, X., *et al.*, PEG precipitation coupled with chromatography is a new and sufficient method for the purification of botulinum neurotoxin type B [corrected]. *PLoS One* 2012, *7*, e39670-e39670.

[26] Li, L., Kantor, A., Warne, N., Application of a PEG precipitation method for solubility screening: a tool for developing high protein concentration formulations. *Protein Sci.* 2013, *22*, 1118-1123.

[27] Hammerschmidt, N., Hobiger, S., Jungbauer, A., Continuous polyethylene glycol precipitation of recombinant antibodies: Sequential precipitation and resolubilization. *Process Biochem* 2016, *51*, 325-332.

[28] Mahadevan, H., Hall, C. K., Statistical-mechanical model of protein precipitation by nonionic polymer. *AIChE J* 1990, *36*, 1517-1528.

[29] Sharon, N., Lis, H., History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* 2004, *14*, 53R-62R.

[30] Cohn, E. J., THE PHYSICAL CHEMISTRY OF THE PROTEINS. *Physiol. Rev.* 1925, *5*, 349-437.

[31] Juckles, I. R. M., Fractionation of proteins and viruses with polyethylene glycol. *Biochim. Biophys. Acta* 1971, *229*, 535-546.

[32] Sim, S. L., He, T., Tscheliessnig, A., Mueller, M., *et al.*, Protein precipitation by polyethylene glycol: a generalized model based on hydrodynamic radius. *J. Biotechnol.* 2012, *157*, 315-319.

[33] Odijk, T., Depletion Theory and the Precipitation of Protein by Polymer. *J. Phys. Chem. B* 2009, *113*, 3941-3946.

[34] Lee, J.-W., Forciniti, D., Effect of glycosylation on the partition behavior of a human antibody in aqueous two-phase systems. *Biotechnol. Prog.* 2013, *29*, 943-950.

[35] Edmond, E., Ogston, A. G., An approach to the study of phase separation in ternary aqueous systems. *Biochem. J.* 1968, *109*, 569-576.

[36] Fee, C. J., Van Alstine, J. M., Prediction of the Viscosity Radius and the Size Exclusion Chromatography Behavior of PEGylated Proteins. *Bioconjug. Chem.* 2004, *15*, 1304-1313.

[37] Kuga, S., Pore size distribution analysis of gel substances by size exclusion chromatography. *J. Chromatogr. A* 1981, *206*, 449-461.

[38] Dohmen, M. P. J., Pereira, A. M., Timmer, J. M. K., Benes, N. E., Keurentjes, J. T. F., Hydrodynamic Radii of Polyethylene Glycols in Different Solvents Determined from Viscosity Measurements. *J. Chem. Eng. Data* 2008, *53*, 63-65.

[39] Sim, S. L., He, T., Tscheliessnig, A., Mueller, M., *et al.*, Branched polyethylene glycol for protein precipitation. *Biotechnol. Bioeng.* 2012, *109*, 736-746.

[40] Lee, J.-W., Forciniti, D., Purification of human antibodies from transgenic corn using aqueous two-phase systems. *Biotechnol. Prog.* 2010, *26*, 159-167.

[41] Hammerling, F., Ladd Effio, C., Andris, S., Kittelmann, J., Hubbuch, J., Investigation and prediction of protein precipitation by polyethylene glycol using quantitative structure-activity relationship models. *J. Biotechnol.* 2017, *241*, 87-97.

[42] Sola, R. J., Griebenow, K., Influence of modulated structural dynamics on the kinetics of α-chymotrypsin catalysis. *FEBS J.* 2006, *273*, 5303-5319.

[43] Sola, R. J., Rodriguez-Martinez, J. A., Griebenow, K., Modulation of protein biophysical properties by chemical glycosylation: biochemical insights and biomedical implications. *Cell. Mol. Life Sci.* 2007, *64*, 2133-2152.

[44] Lins, L., Thomas, A., Brasseur, R., Analysis of accessible surface of residues in proteins. *Protein Sci.* 2003, *12*, 1406-1417.

[45] Miekka, S. I., Ingham, K. C., Influence of hetero-association on the precipitation of proteins by poly(ethylene glycol). *Arch. Biochem. Biophys.* 1980, *203*, 630-641.

[46] Zhu-Salzman, K., Shade, R. E., Koiwa, H., Salzman, R. A., *et al.*, Carbohydrate binding and resistance to proteolysis control insecticidal activity of Griffonia simplicifolia lectin II. *Proc. Natl. Acad. Sci. U.S.A.* 1998, *95*, 15123-15128.

II. SOLUBILITY OF LANOSTEROL IN ORGANIC SOLVENTS AND IN WATER-ALCOHOL MIXTURES AT 101.8 KPA

ABSTRACT

Lanosterol is a sterol derivative whose physicochemical properties are poorly understood. Pure lanosterol (>95%) was isolated from a crude product (54.6%) by a newly developed C18 reverse-phase high-performance liquid chromatography (HPLC) method. Purity and structure were confirmed by gas chromatography-mass spectrometry (GC-MS). The melting temperature and fusion enthalpy were determined to be 408.27 K and 23.61 $kJ \cdot mol^{-1}$, respectively. The solubility of lanosterol was measured in methanol, ethanol, acetonitrile, acetone, N,N-dimethylformamide (DMF), dimethyl sulfoxide (DMSO), ethyl acetate, isopropanol, *n*-propanol, water and binary system, water-methanol, water-isopropanol and water-ethanol using a static equilibrium set up from 278.09 K to 338.78 K. The solubility of lanosterol increases with an increase in temperature. The mole fraction solubility of lanosterol in organic solvents has a minimum of 3.00×10^{-5} in methanol at 277.78 K and a maximum of 0.0048 in *n*-propanol at 318.93 K. The solubility of lanosterol in organic solvents and in water-alcohol mixtures were correlated by the modified Apelblat equation and by the Wilson, NRTL, and UNIQUAC models. In addition, the binary water-alcohol systems were correlated with Apelblat–Jouyban-Acree model and van't Hoff–Jouyban-Acree.

1. INTRODUCTION

Lanosterol was first discovered in the non-saponifiable portion of lanolin. It is a tetracyclic triterpenoid that is synthesized in plants, animals and yeast.[1] It is also an intermediate in the biosynthesis of cholesterol.[2] Lanosterol inhibits the formation of paraneoplastic lesions in the colon of rat.[3] A side chain derivative of lanosterol (3β-hydroxy-5α-lanosta-8, 24-diene) acts as an inhibitor of Δ24 (25) sterol methyl transferase.[4] Lanosterol demethylase, 14-α-demethylase, is the primary target of antifungal drug.[5-7]In addition, besides its biochemical activity, lanosterol is the starting material of other steroids.[8]

The traditional isolation methods of lanosterol involve toxic or hazardous reagents like mercury (II) acetate and Li Al H_4.[9] Even though an environment-friendly routine has been reported recently,[10] it also requires a relative long isolation process and experienced operators.

Recently, lanosterol has been found to dissolve protein aggregates in cataracts.[11] Unfortunately, the restoration of the affected lens only happened in animals (dogs and rabbits).[11] It has been argued that its ineffectiveness to reverse human cataracts is caused by the limited solubility of lanosterol in the delivery medium.[12] One means to increase the solubility of lanosterol is to use mixed-solvents. To the best of our knowledge, the solubility of lanosterol has not been previously reported in the open literature. We did find a product data sheet from Cayman Chemical Company.[13] They reported a solubility of lanosterol of 0.25 mg/ml and 3 mg/ml in ethanol and DMF respectively. The temperature was not specified.

In this report we present a new, fast and environmentally friendly purification process for lanosterol. This new method would make large amounts of highly pure sterol readily available. The solubility of the purified product was experimentally studied in acetone, acetonitrile, N,N-dimethylformamide (DMF), dimethyl sulfoxide (DMSO), ethyl acetate, ethanol, methanol, isopropanol, *n*-propanol and binary water-methanol, water-isopropanol and water-ethanol mixtures as a function of temperature at 101.8 kPa.

2. EXPERIMENTAL METHODS

2.1. MATERIALS

Crude lanosterol with a purity of 54.6% was purchased from Steraloids Inc. (Newport, IR). Crude lanosterol was purified with a Pre-RP-HPLC column and its purity was confirmed by gas-chromatography-mass-spectroscopy (GC-MS). Ultrafiltered type 1 water (≥18.0 MOhm/cm) was utilized as obtained from a Nanopure water system (Barnstead). Acetone, acetonitrile, DMF, DMSO, ethyl acetate, ethanol, isopropanol, *n*-propanol and methanol were ACS grade or better. All solvents were used "as received" without further purification. Details of solvents were summarized in Table 1.

Table 1. Sources and Mass Fraction Purity of Materials

Chemical name	CAS number	Source	% purity	Purification method	Analysis method	Solvent group
lanosta-8,24-dien-3-ol (Lanosterol)	79-63-0	Steraloids Inc.	≥95.0	prep-RF-HPLC[d]	GC[b]	
Methanol	67-56-1	Alfa Aesar	≥99.8	none	GC	protic

Table 1. Sources and Mass Fraction Purity of Materials (cont.)

Chemical name	CAS number	Source	% purity	Purification method	Analysis method	Solvent group
Ethanol (200 proof)	64-17-5	Aldrich	≥99.5	none	GC	protic
Propan-2-ol (Isopropanol)	67-63-0	Aldrich	≥99.5	none	GC	protic
Propan-1-ol(*n*-Propanol)	71-23-8	Aldrich	≥99.7	none	GC	protic
Propan-2-one (Acetone)	67-64-1	Fisher Scientific	99.8	none	GC	aprotic
Acetonitrile	75-05-8	Fisher Scientific	99.9	none	GC	aprotic
N,N-Dimethylmethanamide (DMF)	68-12-2	Fisher Scientific	99.9	none	GC	aprotic
Dimethyl sulfoxide (DMSO)	67-68-5	Fisher Scientific	99.9	none	GC- FID[c]	aprotic
Ethyl Acetate	141-78-6	Fisher Chemical	≥99.5	none	GC	aprotic

[a] Provided by the suppliers.

[b] Gas chromatography.

[c] Gas chromatography with flame ionization detector.

[d] Preparative reverse-phase high-performance liquid chromatography.

2.2. EXPERIMENTAL APPARATUS FOR SOLUBILITY DETERMINATION

The apparatus consists of a 25 mL three-neck glass flask with a water jacket to maintain the temperature. An Amporbe TMD-52 K-type thermocouple (Everett, WA) with a standard uncertainty of 0.01 K was used to monitor the solvent's temperature in the glass flask. In order to keep atmospheric pressure and avoid solvent evaporation, a condenser was coupled to the glass flask. Magnetic stirrers were used in the water jacket and the flask to keep homogeneous temperature and lanosterol concentration. The

temperature in the water jacket was maintained by a Polyscience 9102 circulating bath (Niles, Illinois) with a temperature stability of 0.01 K. The schematic diagram of the experimental set up is shown in Figure S1.

2.3. ISOLATION OF LANOSTEROL

Crude lanosterol is a light yellow powder with a purity of 54.6% according to the manufacturer. The reverse phase chromatographic separation was done in an AKTA purifier system with an Agilent Eclipse XDB-C18 column (4.6mm×150mm). The mobile phase was methanol at a flow rate of 2 mL/min, the injection loop was 1mL, and the UV detector was set at 215nm. The lanosterol and dihydrolanosterol fractions were collected by a fraction collector. Fractions were pooled and vacuum-oven-dried before stored in a desiccator. Pure lanosterol and dihydrolanosterol are white powders.

The purity and identification of lanosterol and dihydrolanoterol were confirmed by an Agilent 7890B gas chromatography system with HP-5MS (Agilent Scientific, USA) capillary column (30m length×0.25mm I.D.0.25mm film thickness) and an Agilent 5977B mass selective detector (MSD). The carrier gas was helium (>99.999%) with a constant flow rate of 1 mL/min. The injection was in split-less-mode at 260 ^{o}C with a 20 min solvent delay and an injection volume of 1μL. The column oven temperature program was set at 150 °C for the first 2 min, ramped to 230 °C (25 °C/min) further on to 248 °C (1.0 °C/min) and finally to 325 °C (25 °C/min) with a 3.7 min hold. The transfer line temperature was set at 300 ^{o}C. The MSD analyzer was set at 70 eV and the electron impact source temperature was 230 ^{o}C.

2.4. THERMAL ANALYSIS

The melting temperature and enthalpy of fusion of lanosterol were determined by a Q2000 Differential Scanning Calorimeter (DSC) (TA instruments). Lanosterol powder was vacuum-oven-dried before taking it to the DSC. About 2 mg of lanosterol was put in a closed DSC pan. An empty DSC pan was used as a blank. The samples were scanned from 338.15 K to 433.15 K with a heating rate of 5.0 K/min. DSC experiments were done in triplicate.

2.5. UV-VIS SPECTROSCOPY

Uv-vis spectra of lanosterol in methanol, the mobile phase in the HPLC experiments, were acquired by a HITACHI U2900 Uv-vis spectrometer. The spectra cover 190 nm to 1100nm and were acquired at room temperature.

2.6. SOLUBILITY DETERMINATION

Five or six mL of solvent and an excess amount of lanosterol were introduced into the 25 mL glass flask described previously. The desired temperature was set and the actual temperature in the glass vessel was recorded. Saturation was reached by adding lanosterol to the solvent until it did not dissolve anymore. Different mixing and settle down times were tested to determine a suitable equilibrium time. A combination of six hours mixing time and 6 hours settle down time was enough to reach equilibrium in all solvents. A sample of the upper clear portion of the mixture was withdrawn with a preheated or precooled glass pipette, transferred to a 1 mL microcentrifuge tube, diluted

100 times or to a proper concentration, and analyzed by HPLC. Each test was done in triplicate.

2.7. HPLC ANALYSIS

The concentration of lanosterol was determined by HPLC. A Shimadzu LC-20AB HPLC system with an Agilent ZORBAX SB-C18 column (2.1×50mm), CTO-10AS column oven, SIL-20AS autosampler and SPD-M20A diode array detector was used. Methanol was used as the mobile phase at a flow rate of 0.2 mL/min. The column oven temperature was 30 °C, the injection volume was 5 μL, and the UV detector wavelength was set at 215 nm.

3. DATA CORRELATION

The temperature dependence of the solubility was correlated by a modified Apelblat equation (Eq.1), which is derived from the van't Hoff isochore by assuming that the apparent partial molar enthalpy of the solute is a linear function of temperature.[14-16]

$$\ln x = A + \frac{B}{T/K} + C\ln(T/K) \tag{1}$$

where x is molar solubility of lanosterol in the solvent; and A, B and C are fitting parameters.

The solubility of a solid at constant pressure can be described by,[17]

$$\ln(\frac{1}{\gamma_i x_i}) = \frac{\Delta_{fus}H_{t,i}}{RT_{t,i}}(\frac{T_{t,i}}{T}-1) - \frac{\Delta C_{p,i}}{R}(\frac{T_{t,i}}{T}-1) + \frac{\Delta C_{p,i}}{R}\ln\frac{T_{t,i}}{T} \tag{2}$$

where x_i is the molar solubility of the solute, γ_i is the activity coefficient of the solute *i* in the liquid phase, R is the gas constant, $\Delta_{fus}H_{t,i}$ is the molar enthalpy of fusion of the solute at the triple point temperature, $T_{t,i}$ is the triple point temperature of the solute, and $\Delta C_{p,i}$ is the differential molar heat capacity of the pure solute. Two assumptions are usually made[17-19] in Eq.2 that only introduce a slight error. The triple point and melting point temperatures are usually close to each other. Therefore, the triple point temperature $T_{t,i}$ can be substituted by the melting temperature $T_{m,i}$. Moreover, the enthalpy of fusion at the triple point temperature $\Delta_{fus}H_{t,i}$ can be substituted by the enthalpy of fusion at the melting point temperature $\Delta_{fus}H_{m,i}$. Second, the last two terms on the right-hand side of Eq.2 are of opposite signs and of similar magnitude; therefore, they cancel each other. Then Eq.2 may be rewritten as,

$$\ln(\frac{1}{\gamma_i x_i}) = \frac{\Delta_{fus}H_{m,i}}{RT_{m,i}}(\frac{T_{m,i}}{T} - 1) \tag{3}$$

3.1. WILSON MODEL

The Wilson's equation[20] provides a good representation of excess Gibbs energies for a variety of miscible mixtures, particularly for solutions of polar components in nonpolar solvents.[17] In a binary liquid-solid system, the activity coefficients are given by,

$$\ln\gamma_1 = -\ln(x_1 + \Lambda_{12}x_2) + x_2(\frac{\Lambda_{12}}{x_1 + \Lambda_{12}x_2} - \frac{\Lambda_{21}}{x_2 + \Lambda_{21}x_1}) \tag{4}$$

$$\ln\gamma_2 = -\ln(x_2 + \Lambda_{21}x_1) - x_1(\frac{\Lambda_{12}}{x_1 + \Lambda_{12}x_2} - \frac{\Lambda_{21}}{x_2 + \Lambda_{21}x_1}) \tag{5}$$

$$\Lambda_{12} = \frac{V_2}{V_1}\exp(-\frac{\lambda_{12}-\lambda_{11}}{RT}) = \frac{V_2}{V_1}\exp(-\frac{\Delta\lambda_{12}}{RT}) \quad (6)$$

$$\Lambda_{21} = \frac{V_1}{V_2}\exp(-\frac{\lambda_{21}-\lambda_{22}}{RT}) = \frac{V_1}{V_2}\exp(-\frac{\Delta\lambda_{21}}{RT}) \quad (7)$$

where Λ_{ij} are two adjustable parameters, λ_{ij} is the energy of interaction between molecules i and j, and V_i is the molar volume of the component i.

3.2. NRTL MODEL

The nonrandom two-liquid equation was first introduced by Renon.[21] Unlike the Wilson model, the NRTL model is applicable to partially miscible as well as completely miscible systems. The activity coefficients a binary mixture are given by,

$$\ln\gamma_1 = x_2^2\left[\tau_{21}(\frac{G_{21}}{x_1+x_2G_{21}})^2 + \frac{\tau_{12}G_{12}}{(x_2+x_1G_{12})^2}\right] \quad (8)$$

$$\ln\gamma_2 = x_1^2\left[\tau_{12}(\frac{G_{12}}{x_2+x_1G_{12}})^2 + \frac{\tau_{21}G_{21}}{(x_1+x_2G_{21})^2}\right] \quad (9)$$

$$\tau_{12} = \frac{g_{12}-g_{22}}{RT} = \frac{\Delta g_{12}}{RT} \quad \tau_{21} = \frac{g_{21}-g_{11}}{RT} = \frac{\Delta g_{21}}{RT} \quad (10)$$

$$G_{12} = \exp(\alpha_{12}\tau_{12}) \quad G_{21} = \exp(\alpha_{12}\tau_{21}) \quad (11)$$

where G_{ij} are adjustable parameters, Δg_{ij} is the energy of interaction between molecules i and j and α_{12} is a nonrandom parameter which varies from 0.20 to 0.47.

3.3. UNIQUAC MODEL

The UNIQUAC equation for excess Gibbs energy g^E consist of two parts, a combinatorial part and a residual part.[23] The activity coefficients for a binary mixture are given by,

$$\ln \gamma_1 = \ln \frac{\Phi_1}{x_1} + \frac{z}{2} q_1 \ln \frac{\theta_1}{\Phi_1} + \Phi_2 (l_1 - \frac{r_1}{r_2} l_2) \\ -q_1 \ln(\theta_1 + \theta_2 \tau_{21}) + \theta_2 q_1 (\frac{\tau_{21}}{\theta_1 + \theta_2 \tau_{21}} - \frac{\tau_{12}}{\theta_2 + \theta_1 \tau_{12}}) \tag{12}$$

$$\ln \gamma_2 = \ln \frac{\Phi_2}{x_2} + \frac{z}{2} q_2 \ln \frac{\theta_2}{\Phi_2} + \Phi_2 (l_2 - \frac{r_2}{r_1} l_1) \\ -q_2 \ln(\theta_2 + \theta_1 \tau_{12}) + \theta_1 q_2 (\frac{\tau_{12}}{\theta_2 + \theta_1 \tau_{12}} - \frac{\tau_{21}}{\theta_1 + \theta_2 \tau_{21}}) \tag{13}$$

$$\Phi_1 = \frac{x_1 r_1}{x_1 r_1 + x_2 r_2} \Phi_2 = \frac{x_2 r_2}{x_1 r_1 + x_2 r_2} \tag{14}$$

$$\theta_1 = \frac{x_1 q_1}{x_1 q_1 + x_2 q_2} \theta_2 = \frac{x_2 q_2}{x_1 q_1 + x_2 q_2} \tag{15}$$

$$l_1 = \frac{z}{2}(r_1 - q_1) - (r_1 - 1) \tag{16}$$

$$l_2 = \frac{z}{2}(r_2 - q_2) - (r_2 - 1) \tag{17}$$

$$\tau_{12} = \exp\left(-\frac{\Delta u_{12}}{RT}\right) \equiv \exp\left(-\frac{a_{12}}{T}\right) \tag{18}$$

$$\tau_{21} = \exp\left(-\frac{\Delta u_{21}}{RT}\right) \equiv \exp\left(-\frac{a_{21}}{T}\right) \tag{19}$$

where τ_{ij} are adjustable parameters, Δu_{ij} are characteristic energies, the coordination number z is set to be 10. The area parameters, q, and volume parameters, r, for lanosterol

and the solvents were calculated according to references,[22-25] and summarized in Table 2.

Table 2. UNIQUAC Parameters of Lanosterol and Solvents

Compound	Molecular weight ($g \cdot mol^{-1}$)	UNIQUAC	
		q	*r*
Lanosterol	426.05	15.4570	19.1675
Methanol	32.04	1.4320	1.4311
Ethanol	46.07	2.5880	2.5755
Isopropanol	60.10	3.1240	3.2491
n-Propanol	60.10	3.1280	3.2499
Acetone	58.08	2.2960	2.5735
Acetonitrile	41.05	1.7240	1.8701
DMF	73.09	2.7360	3.0856
DMSO	78.13	2.4720	2.8266
Ethyl Acetate	88.11	3.1160	3.4786

A linear temperature dependence of the adjustable parameters of the three models described above was assumed.[26] Therefore, Λ_{ij} in the Wilson model, τ_{ij} in the NRTL model and the τ_{ij} in the UNIQUAC model are given by Eq. 20 to 22.

$$\Lambda_{ij} = \frac{V_j}{V_i} \exp\left(a_{ij} + \frac{b_{ij}}{(T/K)} \right) \tag{20}$$

$$\tau_{ij} = a_{ij} + \frac{b_{ij}}{T/K} \tag{21}$$

$$\tau_{ij} = \exp\left(a_{ij} + \frac{b_{ij}}{(T/K)} \right) \tag{22}$$

where a_{ij} and b_{ij} are fitting parameters that are independent of composition and temperature.

3.4. APELBLAT–JOUYBAN-ACREE MODEL AND VAN'T HOFF–JOUYBAN-ACREE

The Apelblat –Jouyban-Acree model (Eq. 23) and van't Hoff–Jouyban-Acree model (Eq. 24) [27] are classic thermodynamic models which used to correlate the binary system.

$$\ln x_3 = w_1\left(A_1 + \frac{B_1}{T/K} + C_1 \ln(T) \right) + w_2\left(A_2 + \frac{B_2}{T/K} + C_2 \ln(T) \right) + \frac{w_1 w_2}{T/K} \sum_{i=0}^{2} J_i \left(w_1 - w_2 \right)^i \tag{23}$$

$$\ln x_3 = w_1\left(A_1 + \frac{B_1}{T/K} \right) + w_2\left(A_2 + \frac{B_2}{T/K} \right) + \frac{w_1 w_2}{T/K} \sum_{i=0}^{2} J_i \left(w_1 - w_2 \right)^i \tag{24}$$

where, x_3 is the mole fraction of lanosterol in binary mixture, w_1 and w_2 are water mass fraction and alcohol mass fraction free of lanosterol. A_1, B_1 C_1 A_2, B_2, C_2, J_0, J_1, *and* J_2 are fitting parameters.

4. RESULTS AND DISCUSSION

Lanosterol purified by Pre-RP-HPLC has a purity of >95% by GC-MS. Additionally, dihydrolanosterol, which is the major impurity in commercial lanosterol, was obtained simultaneously. Lanosterol and dihydrolanosterol were identified using the

standard mass spectra of National Institute of Standards and Technology (NIST) MS spectral library and relevant references.[28] Ion groups at m/z of corresponding steroid were monitored, lanosterol (69,109,393,411,426) and dihydrolanosterol (43,69,395,413,428). The chromatogram and MS scan are shown in Figure S2 and S3.

The solubility of lanosterol in methanol, ethanol, isopropanol, *n*-propanol, acetonitrile, acetone, DMF, DMSO, and ethyl acetate as a function of temperature is summarized in Table 3 and is plotted in Figure 1. The solubility of lanosterol in water−methanol, water−ethanol, and water−isopropanol mixtures is summarized in Table 4 and is plotted in Figures 2−4. The van't Hoff plots of organic solvents and water−methanol, water−ethanol, and water−isopropanol mixtures are shown in Figures S4−S7.

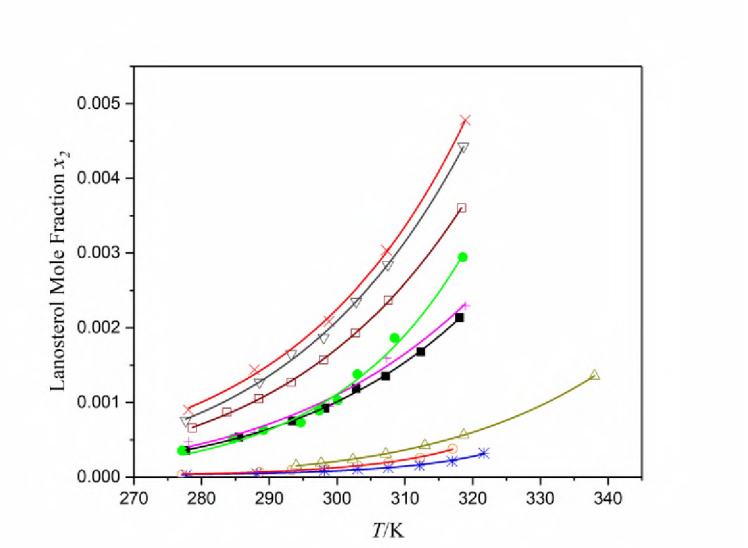

Figure 1. The mole fraction of lanosterol x_2 in selected solvents at different temperatures.▽, ethyl acetate; □, acetone; ■, ethanol; ●, DMF; △, DMSO; ○, acetonitrile; ✳, methanol; +,isopropanol; ×, *n*-propanol. The solid line are fittings of data by the modified Apelblat equation

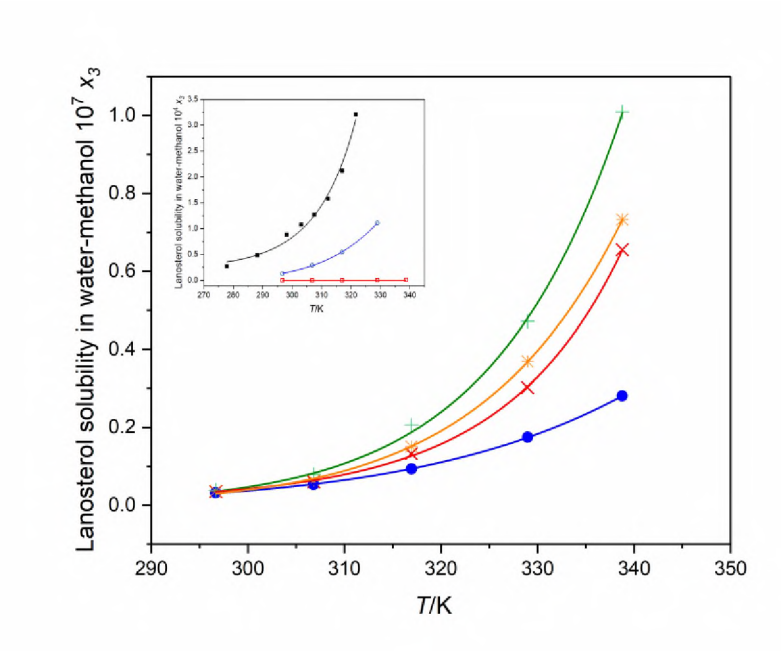

Figure 2. The mole fraction of lanosterol (3) x_3 in water and water (1)-methanol (2) mixture at different temperatures. ●, water v_2=0.00; ×, v_2=0.050; ✳, v_2=0.20; ＋, v_2=0.50; The insert shows the x_3 in □, v_2=0.70; ○, v_2=0.90 and ■ v_2=1.00 ; —, the solid line are fittings of data by the modified Apelblat equation

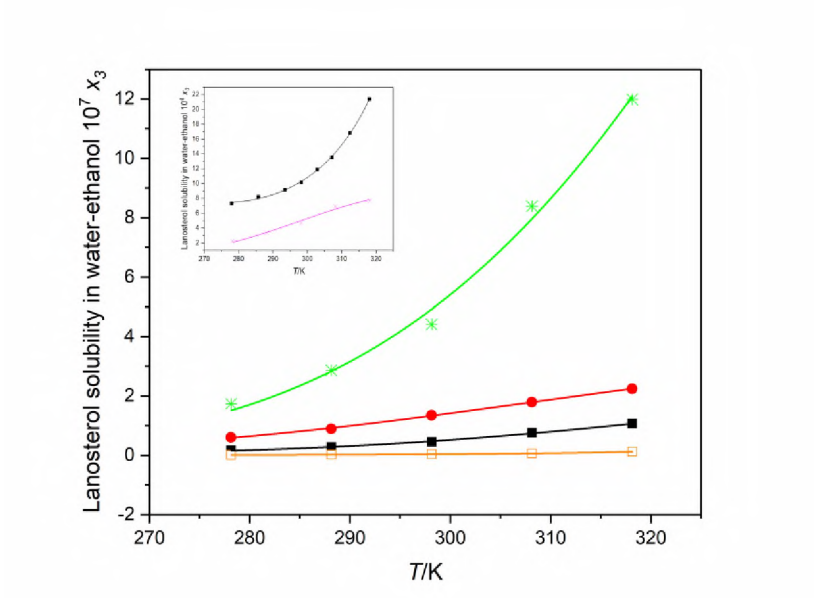

Figure 3. The mole fraction of lanosterol (3) x_3 in water and water (1)-ethanol (2) mixture at different temperatures. □, water v_2=0.00; ■, v_2=0.050; ●, v_2=0.30; ✳, v_2=0.60; the insert shows the ×, v_2=0.90; ○, v_2=1.00, —, the solid line are fittings of data by the modified Apelblat equation

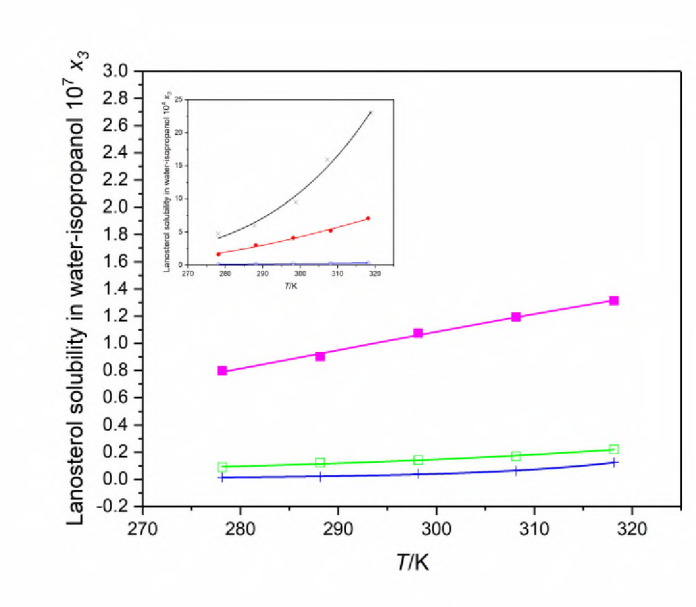

Figure 4. The mole fraction of lanosterol (3) x_3 in water and water (1)-isopropanol (2) mixture at different temperatures.＋, water v_2=0.00; □, v_2=0.050; ■, v_2=0.30; the insert shows the ○, v_2=0.60; ●, v_2=0.90; ×, v_2=1.00; —, the solid line are fittings of data by the modified Apelblat equation

Table 3. Mole fraction solubility x_2 of lanosterol in organic solvents at temperature range (277.09K-337.99K) under 101.8 KPa[a]

T/K	1000x_2	*T*/K	1000x_2	*T*/K	1000x_2
Methanol (A)[b]		Ethanol (D)[b]		Isopropanol (D)[b]	
277.78	0.03	277.95	0.36	278.07	0.47
288.15	0.05	285.74	0.53	287.81	0.60
298.15	0.09	293.45	0.75	298.84	0.95
303.05	0.11	298.25	0.92	307.27	1.59
307.55	0.13	302.85	1.19	318.93	2.30
312.15	0.16	307.15	1.35		

Table 3. Mole fraction solubility x_2 of lanosterol in organic solvents at temperature range (277.09K-337.99K) under 101.8 KPa[a] (cont.)

T/K	$1000x_2$	*T*/K	$1000x_2$	*T*/K	$1000x_2$
Methanol (A)[b]		Ethanol (D)[b]		Isopropanol (D)[b]	
316.95	0.21	312.35	1.68		
321.65	0.32	318.05	2.14		
n-propanol (D)[b]		Acetone (A)[b]		Acetonitrile (A)[b]	
278.07	0.90	278.65	0.66	277.09	0.03
287.81	1.44	283.75	0.87	288.55	0.06
298.84	2.09	288.45	1.05	293.35	0.09
307.27	3.04	293.25	1.27	298.15	0.13
318.93	4.78	298.05	1.57	303.05	0.15
		302.75	1.93	307.55	0.19
		307.55	2.37	312.25	0.26
		318.36	3.60	317.05	0.38
DMF (amorphous)[b]		DMSO (amorphous)[b]		Ethyl Acetate (A)[b]	
277.15	0.35	293.95	0.15	277.67	0.76
284.65	0.51	297.65	0.19	288.55	1.27
289.15	0.63	302.35	0.24	293.25	1.65
294.65	0.73	307.15	0.31	298.05	1.87
297.35	0.89	312.95	0.42	302.85	2.35
300.15	1.03	318.69	0.57	307.45	2.84
303.05	1.38	337.99	1.36	318.64	4.43
308.46	1.87				

Table 3. Mole fraction solubility x_2 of lanosterol in organic solvents at temperature range (277.09K-337.99K) under 101.8 KPa[a] (cont.)

T/K	$1000x_2$	T/K	$1000x_2$	T/K	$1000x_2$
DMF (amorphous)[b]		DMSO (amorphous)[b]		Ethyl Acetate (A)[b]	
318.55	2.94				

[a]x_2 is the experimental mole fraction solubility of lanosterol at temperature T; the standard uncertainties u are $u(T)$ = 0.01 K, $u(P)$ = 0.22 kPa; the relative standard uncertainty ur is $ur(x_2)$ = 0.023.

[b]Lanosterol crystal forms in equilibrium with solvents.

Table 4. Mole fraction solubility x_3 of lanosterol (3) in water, water(1)-methanol(2), water(1)-ethanol(2) and water(1)-isopropanol (2) binary mixtures at the temperature range (278.15 K-338.78 K) under 101.8KPa[a]

T/K	10^6x_3					
	water (1)- methanol(2) (A)[b]					
	water	v_2=0.050, w_2=0.040	v_2=0.20, w_2=0.16	v_2=0.50, w_2=0.44	v_2=0.70, w_2=0.65	v_2=0.90, w_2=0.88
296.67	0.0032	0.0035	0.0037	0.0038	0.0648	13.0
306.81	0.0053	0.0059	0.0063	0.0079	0.116	29
316.94	0.0093	0.0133	0.0151	0.0206	0.223	54
328.97	0.0175	0.030	0.0370	0.047	0.47	111
338.78	0.028	0.067	0.073	0.101	1.15	
	water (1)- ethanol(2) (A)[b]					
	water	v_2=0.050, w_2=0.040	v_2=0.30, w_2=0.25	v_2=0.60, w_2=0.54	v_2=0.90, w_2=0.88	
278.15	0.00141	0.0168	0.061	0.174	223	

Table 4. Mole fraction solubility x_3 of lanosterol (3) in water, water(1)-methanol(2), water(1)-ethanol(2) and water(1)-isopropanol (2) binary mixtures at the temperature range (278.15 K-338.78 K) under 101.8KPa[a] (cont.)

T/K	$10^6 x_3$				
	water (1)- ethanol(2) (A)[b]				
	water	v_2=0.050, w_2=0.040	v_2=0.30, w_2=0.25	v_2=0.60, w_2=0.54	v_2=0.90, w_2=0.88
288.15	0.00193	0.028	0.090	0.29	334
298.15	0.0042	0.046	0.141	0.44	468
308.15	0.0062	0.076	0.178	0.84	691
318.15	0.0123	0.114	0.221	1.20	769
	water (1)- isopropanol(2) (A)[b]				
	water	v_2=0.050, w_2=0.040	v_2=0.30, w_2=0.25	v_2=0.60, w_2=0.54	v_2=0.90, w_2=0.88
278.15	0.00141	0.0088	0.080	10.70	162
288.15	0.00193	0.0122	0.090	14.2	300
298.15	0.0042	0.0139	0.113	20.8	412
308.15	0.0062	0.0171	0.123	25	522
318.15	0.0123	0.0223	0.130	31	708

[a]x_3 is the experimental mole fraction solubility of lanosterol at temperature T, v_2 is volume fraction of alcohols in water-alcohols binary system free of lanosterol, w_2 is mass fraction of alcohols in water-alcohols binary system free of lanosterol; the standard uncertainties u are $u(T)$ = 0.01 K, $u(P)$ = 0.22 kPa; the relative standard uncertainty ur is $ur(x_3)$ = 0.043.

[b]Lanosterol crystal forms in equilibrium with solvents.

As shown in Figure 1, the solubility of lanosterol increases with increasing temperature in all the solvents studied. From 277.78 K to 338.15 K, the solubility of lanosterol is the largest in *n*-propanol and the lowest in methanol. The solubility of lanosterol in DMF is lower than in ethanol below 300.15 K but it is higher above 300.15 K. Furthermore, below 298.15 K the lanosterol solubility in isopropanol is lower than in the ethanol, above 298.15 K the order between isopropanol and ethanol switched. Around 277.15 K the lanosterol solubility increases according to the following order: methanol< acetonitrile< DMF< isopropanol< acetone< ethanol< ethyl acetate< *n*-propanol and around 318.15 K, the order change to: methanol< acetonitrile< DMSO< ethanol< isopropanol< DMF< acetone< ethyl acetate< *n*-propanol.

The solvents were sorted into two groups according to their H-bond capacity. The protic solvents group includes methanol, ethanol, *n*-propanol and isopropanol, and the aprotic solvents group includes DMSO, DMF, acetone, acetonitrile and ethyl acetate. In each group, the solubility of lanosterol increases with decreasing solvent polarity except isopropanol. In addition, the solubility of lanosterol increases with decreasing Hanson solubility parameters within each group except isopropanol in the protic solvents group, except acetonitrile in aprotic solvents group. The Hanson solubility parameters, polarity[29, 30] and polarizability[31] of solvents are summarized in Table S1.

Lanosterol has a highly hydrophobic steroid domain with a hydroxyl motif, which makes the molecule weakly polar. The predominant intermolecular forces in the protic solvents group are H-bond and dipole-dipole interactions. Polar protic solvents cannot dissolve lanosterol efficiently because the weak H-bond between lanosterol and ethanol or methanol cannot replace the strong H-bond in pure solvents without an energy penalty.

Lanosterol is less soluble in methanol than in ethanol, isopropanol and *n*-propanol, which is expected because methanol is a better H-bond molecule than other alcohols studied in this work. The solubility of lanosterol increase while increasing alkyl chain length of alcohols except isopropanol, maybe due to the hydroxyl group is in the middle of the isopropanol alkyl chain. The solvents in the aprotic solvents group are only H-bond acceptors. In such systems, the H-bond interactions are weaker than in the protic solvents group. The dipole-dipole interaction in pure solvents become weaker as the polarity decreases, then the weak dipole-dipole interaction between lanosterol and solvents may easily replace the dipole-dipole interaction in pure solvents, causing higher solubility. Dispersion forces or induced-dipole interactions may also contribute to dissolution. Dispersion forces depend on the polarizability of solvents. In the aprotic solvents group, the polarizabilities of each solvent are in the order of ethyl acetate>DMSO>DMF>acetone>acetonitrile, but the solubilities of lanosterol do not follow the order of polarizabilities except for ethyl acetate and acetonitrile. The above observations lead to the conclusion that the lower the polarity of solvent the higher the solubility of lanosterol, and that dispersion forces may also affect the solubility to some degree in some solvents. To increase the solubility of lanosterol, n-hexane (polarity 0.9) was tested. However, the solubility of lanosterol in *n*-hexane (data not shown) is even lower than in ethyl acetate (polarity 23.0). Therefore, the dissolution of lanosterol is not only affected by the polarity of solvent but by a synergy of dipole-dipole interactions, H-bond interactions and dispersion forces. This suggests that a low polarity solvent and a weak H-bond capacity solvent mixtures may enhance the solubility of lanosterol. For

example, it has been found that mixtures of ethanol and hexane increase the solubility of cholesterol.[32]

The addition of alcohols to water increased lanosterol's solubility. That indicates that the alcohols are cosolvents of lanosterol in water. Lanosterol solubility increased with increasing of temperature in water-alcohol mixtures and increased with increasing alcohol content.

The PXRD pattern (Figure S8) of lanosterol before the solubility experiments were conducted is identical to a previously reported pattern[33]. Ref 33 reported PXRD for lanosterol powder and lanosterol crystallized in diisooctylphthalate (DIOP). In the patent owned by Pan *et al.*[34] showed PXRD pattern of different lanosterol crystal forms. According to this patent the crystal form of powder lanosterol is anhydrous Form A. Figure S8 shows that the PXRD patterns changed after solubility experiments in ethanol, isopropanol and *n*-propanol, which indicate that there is another polymorph (Form D, according to the same patent) formed during the experiments. Furthermore, the PXRD experiments show that the solid at equilibrium with solutions of lanosterol in DMSO and DMF has lost the crystalline structure and forms amorphous precipitates. The PXRD also shows that after vacuum-oven drying, the amorphous precipitate obtained from precipitation in DMF and DMSO forms crystalline structures. In the other pure solvents and in water-alcohol mixtures, the PXRD patterns are identical to the one obtained with powder lanosterol. The crystal forms of lanosterol in equilibrium with solvents were submitted in Table 3 and Table 4.

The melting temperature and enthalpy of fusion of the purified lanosterol were determined to be 408.27 K with an uncertainty $u(T_m) = 0.5$ K and 23.61 kJ·mol^{-1} with an

uncertainty $u(\Delta H_{fus}) = 0.14$ kJ·mol^{-1}, respectively. The molar extinction coefficient in methanol at 202 nm was 382,801(M·cm)$^{-1}$. Literature values of melting temperatures are reported as the width of the endothermic peak, the onset of the endothermic peak or its mean value. The melting temperature reported in this work is the onset temperature of the endothermic peak because this value is less sensitive to heating rates and sample mass. It is lower than the values reported by Boar et al.[8] (413.15 K), Maienthal and Franklin[35] (411.15 K-413.15 K), Johnston and Bloch[9] (410.15 K-411.15 K) and Jagodzinski and Rodewald[10] (419.15 K-420.15 K). The differences may be caused by differences in purity, by the presence of different impurities because of the different separation methods used and/or by the use of different experimental protocols. For example, it has been found that the heating rate,[36] which is not reported in some of the data, affects melting temperature measurements. [36] The DSC thermograph is shown in Figure S9.

Figure S10 shows the DSC data for lanosterol solids obtained in solvents whose PXRD patterns are different from the starting material's one. The DSC results for ethanol, isopropanol and *n*-propanol show two endothermic peaks. The peak around 370 K may be caused by eutectic process. The melting temperature and enthalpy of fusion of lanosterol crystal formed in ethanol, isopropanol and *n*-propanol are 390.52 K with an uncertainty u(Tm) = 0.5 K and 6.97 kJ·mol^{-1} with an uncertainty $u(\Delta H_{fus}) = 0.21$kJ·mol^{-1}, respectively. The DSC results of the dried precipitates from DMSO and DMF both show an endothermic peak around 386 K (Figure S10) which agree with PXRD patterns.

The solubility of lanosterol in organic solvents was correlated by the modified Apelblat equation (Table 5), the Wilson, NRTL and UNIQUAC models (Table 6 except

DMSO and DMF). The root mean square deviation (*RMSD*) of the experimental solubility $x_2^{\exp}$ and calculated solubility x_2^{cal} were calculated as follow,

$$RMSD = \left[\frac{\sum_{i=2}^{N} x_i^{cal} - x_i^{\exp}}{N} \right]^{1/2} \tag{25}$$

where N is the number of experimental data in each solvent at a given temperature.

In addition, in order to evaluate each model, the average absolute deviation percentages (AADP) and relative average deviation (RAD) were calculated.

$$AADP = \frac{1}{N} \sum_{N} \left| \frac{x_i^{\exp} - x_i^{cal}}{x_i^{\exp}} \right| 100\% \tag{26}$$

$$RAD = \frac{1}{N} \sum \left(\frac{\left| x_3^{cal} - x_3^{\exp} \right|}{x_3^{\exp}} \right) \tag{27}$$

The molar volume of lanosterol and the solvents were taken from Advanced Chemistry Development ACD/Chemsketch Software (© 1994− 2018 ACD/Laboratories) whereas the melting temperature ($T_{m,i}$) and enthalpy of fusion ($\Delta_{fus}H_{m,i}$) of lanosterol were measured in this work.

Table 5 and Table 6 show that the correlated solubilities by all four models are in good agreement with experimental data. The largest *RMSD* is 9.45×10^{-05} for the Wilson model in *n*-propanol. The largest *AADP* is 8.26% for the NRTL model in isopropanol. Assuming that $\Delta C_{p,i} \sim 0$ may have contributed to the deviation.[32]

Table 5. Parameters of the Apelblat Equation and Root Mean Square Deviation (*RMSD*) and the average absolute deviation percentage (*AADP*) for Lanosterol in Selected Solvents

Solvents	*A*	*B*/T	*C*	10^5 *RMSD*	*AADP* %
Methanol	-823.88	32557.15	123.77	0.92	9.05
Ethanol	-25.90	-2464.60	4.77	2.39	1.22
n-Propanol	-139.54	2899.70	21.70	4.33	2.72
Isopropanol	16.99	-4281.88	-1.67	7.66	8.07
Acetone	-40.66	-1686.88	6.99	1.60	1.13
Acetonitrile	-604.28	22145.50	91.42	0.80	7.42
DMF	88.49	-8649.48	-11.65	7.11	7.58
DMSO	39.58	-6431.80	-4.66	0.20	0.65
Ethyl Acetate	-44.93	-1478.80	7.66	3.85	1.65

The solubility of lanosterol in binary water-alcohol mixtures was correlated by the Apelblat–Jouyban-Acree and the van't Hoff–Jouyban-Acree models. The RMSD, RAD and fitted lanosterol mole fraction x_3^{cal} are shown in Table S2 and Figure S11-S13. The van't Hoff-Jouyban-Acree model fits the data better than the Apelblat-Jouyban-Acree model. Most of the deviations are at the very low solubilities observed at very low organic solvent mole fractions.

The parameters b_{ij}/T in the Wilson model are negligible when compared to a_{ij} (Table 6) which are approximately equal to the $-\frac{\Delta\lambda_{ij}}{RT}$ terms in Eq. 6 and 7. The $\Delta\lambda_{ij}$ in Wilson model of the lanosterol-methanol system are $\Delta\lambda_{12}$=-2932.32 and $\Delta\lambda_{21}$ =13749.52

which are of the same order of magnitude of a similar steroid molecule, desmosterol in methanol ($\Delta\lambda_{12}$=-3827.52 and $\Delta\lambda_{21}$=14298.65).

Table 6. Interaction Parameters and Root Mean Square Deviation (RMSD) and the Average Absolute Deviation Percentage (AADP) for Wilson, NRTL and UNIQUAC Models

Solvent	Wilson					
	a_{12}	b_{12}/K	a_{21}	b_{21}/K	10^5 *RMSD*	*AADP*%
Methanol	2.88	-6.74	-10.37	5.73	0.63	4.61
Ethanol	1.31	0.091	-9.30	4.44	5.15	2.84
Isopropanol	0.97	-0.24	-8.77	4.13	3.97	1.84
n-Propanol	1.00	0.46	-8.09	3.41	9.45	3.04
Acetone	2.07	-2.57	-6.29	3.85	4.97	2.07
Acetonitrile	2.76	-11.40	-10.15	5.15	1.28	6.32
Ethyl Acetate	1.82	-0.93	-5.85	-0.28	7.83	2.81
Solvent	NRTL[a]					
	a_{12}	b_{12}/K	a_{21}	b_{21}/K	10^5 *RMSD*	*AADP*%
Methanol	-4.62	-44.01	3.28	396.66	1.97	5.37
Ethanol	-2.11	-1128.88	2.01	813.17	2.70	1.91
Isopropanol	-4.42	130.31	2.39	575.61	8.62	8.26
n-Propanol	-10.31	2476.33	3.57	23.18	5.74	2.03
Acetone	-3.00	43.04	2.62	235.76	1.65	1.05
Acetonitrile	-1.78	2456.85	-336.24	-364.61	1.36	4.55
Ethyl Acetate	-1.52	-89.38	1.26	295.20	3.77	1.40

Table 6. Interaction Parameters and Root Mean Square Deviation (RMSD) and the Average Absolute Deviation Percentage (AADP) for Wilson, NRTL and UNIQUAC Models (cont.)

Solvent	UNIQUAC					
	a_{12}	b_{12}/K	a_{21}	b_{21}/K	10^5 *RMSD*	*AADP*%
Methanol	-1.76	-125.06	0.71	25.41	0.52	3.46
Ethanol	0.83	-200.20	-1.41	40.28	1.81	1.27
Isopropanol	-1.18	-130.38	0.42	29.89	7.56	3.98
n-Propanol	0.19	-150.04	-0.86	10.89	3.79	1.87
Acetone	0.23	-49.08	-0.90	6.49	1.45	1.01
Acetonitrile	-1.49	-63.40	0.74	-33.07	0.66	3.56
Ethyl Acetate	0.44	-52.48	-1.05	-5.14	3.76	1.38

[a] The nonrandom parameter in the NRTL model is $a = 0.2$

Furthermore, a_{21} increases with decreasing solvent polarity within each solvent groups. This indicates that with decreasing solvent polarity, the interaction energy difference between the solvent and lanosterol and between the pure solvents ($\Delta\lambda_{ij}$) becomes smaller. Then it leads to activity coefficients approximately equal to one. If the activity coefficients are closer to one, the interactions between solute and solvent are similar to solvent-solvent interactions. Therefore, replacement of solvent-solvent bonds by solvent-solute bonds is favorable and the solubility increases. This dependence of the coefficients of Wison model with solvent polarity was not found in the other two models.

The interaction parameters, $\Delta g_{12}=g_{12}-g_{11}$, $\Delta g_{21}=g_{21}-g_{22}$ and $\Delta u_{12}=u_{12}-u_{11}$, $\Delta u_{21}=u_{21}-u_{22}$ for the NRTL and UNIQUAC models are presented in Table S3. The g_{ii} and

u_{ii} are the energy of evaporation of lanosterol in each solvent. The u_{ij} and g_{ij} are the interaction energies between lanosterol and solvents such as g_{12}=g_{21}, u_{12}=u_{21}. In the NRTL model, the Δg_{12} are positive for all solvents, which indicates that the interaction energy between solvent and solute are larger than the evaporation energy of the solvents. Δg_{21} are negative in all solvents. Δu_{12} in the UNIQUAC model are all positive whereas the Δu_{21} are negative for all solvents. Although the Δu_{ij} and Δg_{ij} in both models are related to the evaporation energy of solute and solvents, the values show a large difference. The discrepancy between the two models is likely to be caused by the absence of an entropic term in the NRTL model.[37, 38]

The mixing properties, $\Delta_{mix}G$, $\Delta_{mix}S$ and $\Delta_{mix}H$ were computed using the Wilson Equation and summarized in Table S4. The mixing enthalpies ($\Delta_{mix}H$) in all solvents are positive except for acetonitrile and ethyl acetate, therefore, the dissolution of lanosterol is an endothermic process. Only for acetonitrile and ethyl acetate, the dissolution is exothermic. The mixing entropies ($\Delta_{mix}S$) are positive for all solvents, which indicates all solvent-lanosterol systems become more disorder and the dissolution is an entropy-driven process. The mixing Gibbs energy ($\Delta_{mix}G$) of all solvents are negative and decrease with increasing temperature. The dissolution of lanosterol in all selected solvents is spontaneous.

In the absence of binary interaction parameters, the solubility of lanosterol in methanol-water systems at 298.15 K were predicted by the modified Wilson model

$$-\ln x_m^{Sat} = 1 - \frac{w_1(1+\ln x_1^{Sat})}{w_1 + w_2\lambda_{12}} - \frac{w_2(1+\ln x_2^{Sat})}{w_1\lambda_{21} + w_2} \tag{28}$$

where x_m^{Sat} is the solubility of the solute in the mixed solvents x_1^{Sat} and x_2^{Sat} are the solute solubility in organic solvent and water at a given temperature; w_1 and w_2 are the mass fractions of organic solvent and water; and λ_{ij} are the energy of interaction parameters. Replacing the λ_{ij} with Abraham solute parameters in Eq. 28 we obtain,

$$-\ln x_m^{Sat} = 1 - \frac{w_1(1+\ln x_1^{Sat})}{w_1 + w_2(J_0 + J_1E + J_2S + J_3A + J_4B + J_5V)} - \frac{w_2(1+\ln x_2^{Sat})}{w_1(J'_0 + J'_1E + J'_2S + J'_3A + J'_4B + J'_5V) + w_2} \quad (29)$$

where J_i and J_i' terms are model constants; *E, S, A, B, V* are Abraham solute parameters, which are the excess molar refraction (*E*), the dipolarity/polarizability of the solute (*S*), the solute's hydrogen-bond acidity and basicity(*A, B*) and the McGowan volume of the solute(*V*). Barzegar-Jalali *et al.*[39] suggested a set of *J* and *J'* that were regressed from 41 drug solubility data sets for methanol-water systems. The trained version of the modified Wilson model for methanol-water systems is,

$$-\ln x_m^{Sat} = 1 - \frac{w_1(1+\ln x_1^{Sat})}{w_1 + w_2(0.626 - 0.622E + 0.339S + 0.246A + 0.218B - 0.096V)} - \frac{w_2(1+\ln x_2^{Sat})}{w_1(1.525 + 0.108E - 0.109S + 0.092A + 0.586B - 0.228V) + w_2} \quad (30)$$

The Abraham solute parameters of lanosterol were calculated by ACD/I-Lab software: $E = 1.41$, $S = 1.67$, $A = 0.31$, $B = 0.9$, $V = 3.8739$.

The solubility of lanosterol in methanol-water mixtures was measured and compared with the solubility predicted by Eq. 30 (Figure 5). The molar solubility of lanosterol in water is 3.2×10^{-9}. Interestingly, the modified Wilson model predicted that the lanosterol's solubility first decreases and then increases with increasing methanol volume fraction.

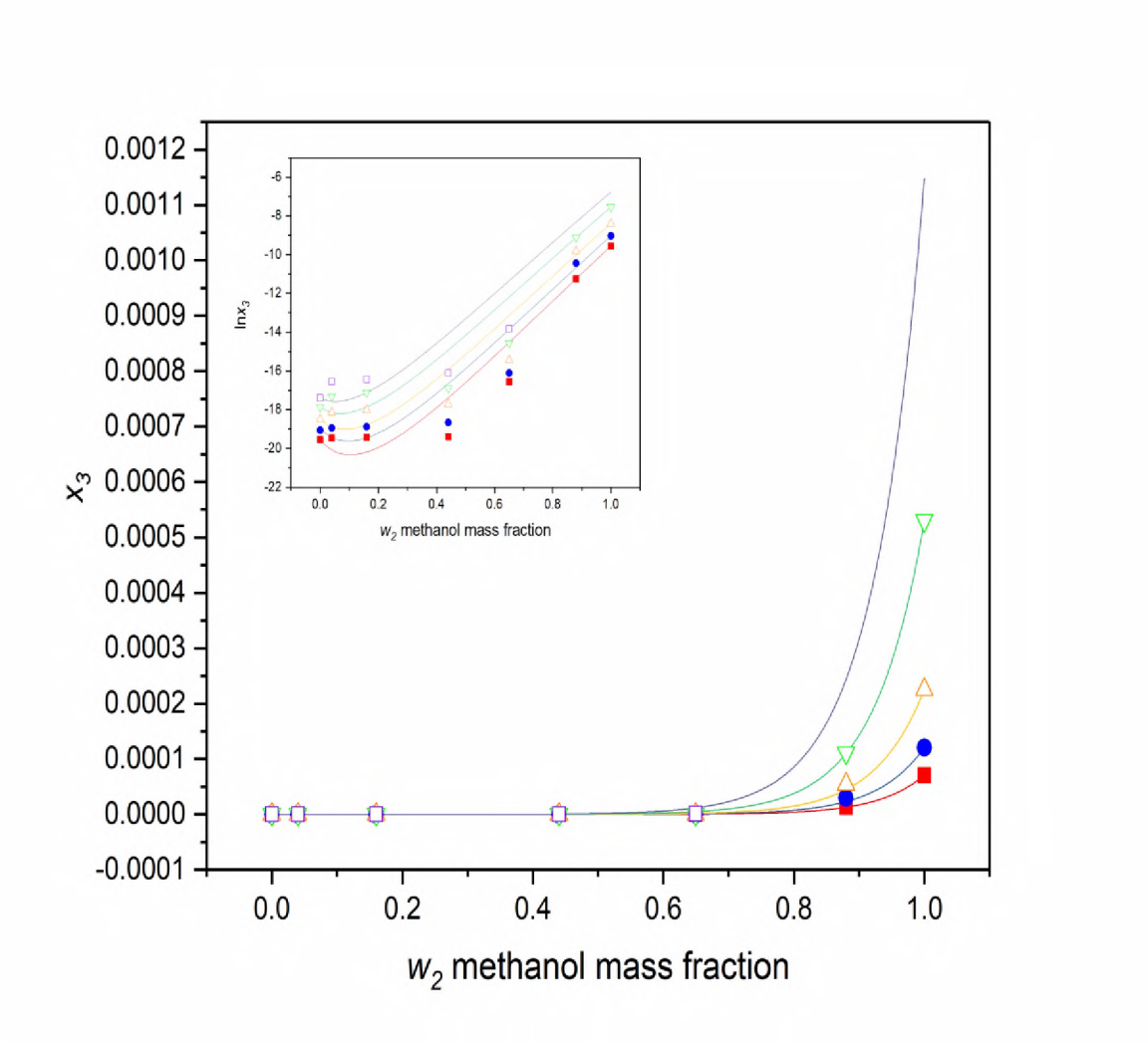

Figure 5. The solubility of lanosterol (x_3) in water (1)-methanol (2) binary mixtures compare with the solubility predicted by modified Wilson model.■, 296.67 K; ●, 306.81 K; △, 316.94 K; ▽, 328.97 K; □, 338.78 K. The solid line represents solubility prediction by the modified Wilson model. The inset shows natural logarithmic solubility

The solubilities predicted by Eq. 30 deviate less from the experimental data at low and high methanol volume fraction. The overall relative deviation is 344 %. The modified Wilson model may only applicable at high methanol content for lanosterol.

5. CONCLUSIONS

The solubility of lanosterol in nine organic solvents from 277.09 K to 338.15 K and in water-alcohol mixtures from 277.78 K to 338.78 K was measured. The melting temperature and enthalpy of fusion of lanosterol were determined to be 408.27 K and 23.61 kJ·mol-1, respectively. The solubility of lanosterol in n-propanol is the largest one among the selected solvents. The activity coefficient models of Wilson, NRTL and UNIQUAC correlated the experimental data satisfactorily. The Apelblat –Jouyban-Acree and the van't Hoff–Jouyban-Acree models correlated water-alcohol binary systems without significant deviations.

REFERENCES

[1] Suzuki, M., Xiang, T., Ohyama, K., Seki, H., *et al.*, Lanosterol Synthase in Dicotyledonous Plants. *Plant and Cell Physiology* 2006, *47*, 565-571.

[2] Chen, N., Wang, S., Smentek, L., Hess, B. A., Wu, R., Biosynthetic Mechanism of Lanosterol: Cyclization. *Angewandte Chemie International Edition* 2015, *54*, 8693-8696.

[3] Rao, C. V., Newmark, H. L., Reddy, B. S., Chemopreventive effect of farnesol and lanosterol on colon carcinogenesis. *Cancer Detection and Prevention* 2002, *26*, 419-425.

[4] Nes, W. D., Sterol methyl transferase: enzymology and inhibition. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* 2000, *1529*, 63-88.

[5] Kushima, H., Tokimatsu, I., Ishii, H., Kawano, R., *et al.*, Cloning of the lanosterol 14-α-demethylase (ERG11) gene in Trichosporon asahii: a possible association between G453R amino acid substitution and azole resistance in T. asahii. *FEMS Yeast Research* 2012, *12*, 662-667.

[6] Koch, A., Kumar, N., Weber, L., Keller, H., *et al.*, Host-induced gene silencing of cytochrome P450 lanosterol C14α-demethylase–encoding genes confers strong resistance to *Fusarium* species. *Proceedings of the National Academy of Sciences* 2013, *110*, 19324-19329.

[7] Sagatova, A. A., Keniya, M. V., Wilson, R. K., Monk, B. C., Tyndall, J. D. A., Structural Insights into Binding of the Antifungal Drug Fluconazole to *Saccharomyces cerevisiae* Lanosterol 14α-Demethylase. *Antimicrobial Agents and Chemotherapy* 2015, *59*, 4982-4989.

[8] Boar, R. B., Lewis, D. A., McGhie, J. F., Isolation and some reactions of lanosterol. A synthesis of agnosterol. *Journal of the Chemical Society, Perkin Transactions 1* 1973, 1583-1588.

[9] Johnston JD, G. F., Bloch K., Isolation of Lanosterol from "Isocholesterol". *J Biol Chem* 1957, *224*, 185-190.

[10] Jagodzinski, J., Rodewald, W., A new method of isolating lanosterol from "isocholesterol". *Polish journal of chemistry* 1978, *52*, 2473-2477.

[11] Zhao, L., Chen, X. J., Zhu, J., Xi, Y. B., *et al.*, Lanosterol reverses protein aggregation in cataracts. *Nature* 2015, *523*, 607-611.

[12] Shanmugam, P., Barigali, A., Kadaskar, J., Borgohain, S., *et al.*, Effect of lanosterol on human cataract nucleus. *Indian Journal of Ophthalmology* 2015, *63*, 888-890.

[13] Chemical, C. Product Information of Lanosterol, 2016.

[14] Grant, D. J. W., Non-linear van't Hoff solubility-temperature plots and their pharmaceutical interpretation. *International journal of pharmaceutics* 1984, *18*, 25-38.

[15] Zorrilla-Veloz, R. I., Stelzer, T., López-Mejías, V., Measurement and Correlation of the Solubility of 5-Fluorouracil in Pure and Binary Solvents. *Journal of Chemical & Engineering Data* 2018, *63*, 3809-3817.

[16] Qiu, J., Song, S., Chen, X., Yi, D., *et al.*, Determination and Correlation of the Solubility of l-Fucose in Four Binary Solvent Systems at the Temperature Range from 288.15 to 308.15 K. *Journal of Chemical & Engineering Data* 2018, *63*, 3760-3768.

[17] Prausnitz, J. M., Lichtenthaler, R. N., Azevedo, E. G. , *Molecular thermodynamics of fluid-phase equilibria*, Prentice Hall PTR., Upper Saddle River, N.J 1999.

[18] Ali, S. H., Al-Mutairi, F. S., Fahim, M. A., Solubility of polycyclic aromatics in binary solvent mixtures using activity coefficient models. *Fluid Phase Equilibria* 2005, *230*, 176-183.

[19] Žilnik, L. F., Jazbinšek, A., Hvala, A., Vrečer, F., Klamt, A., Solubility of sodium diclofenac in different solvents. *Fluid Phase Equilibria* 2007, *261*, 140-145.

[20] Wilson, G. M., Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing. *Journal of the American Chemical Society* 1964, *86*, 127-130.

[21] Renon, H., Prausnitz, J. M., Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE Journal* 1968, *14*, 135-144.

[22] Hansen, H. K., Rasmussen, P., Fredenslund, A., Schiller, M., Gmehling, J., Vapor-liquid equilibria by UNIFAC group contribution. 5. Revision and extension. *Industrial & Engineering Chemistry Research* 1991, *30*, 2352-2355.

[23] Wittig, R., Lohmann, J., Gmehling, J., Vapor−Liquid Equilibria by UNIFAC Group Contribution. 6. Revision and Extension. *Industrial & Engineering Chemistry Research* 2003, *42*, 183-188.

[24] Balslev, K., Abildskov, J., UNIFAC Parameters for Four New Groups. *Industrial & Engineering Chemistry Research* 2002, *41*, 2047-2057.

[25] Peng, C., Chan, M. N., Chan, C. K., The Hygroscopic Properties of Dicarboxylic and Multifunctional Acids: Measurements and UNIFAC Predictions. *Environmental Science & Technology* 2001, *35*, 4495-4501.

[26] Chen, J., Chen, G., Cheng, C., Cong, Y., *et al.*, Thermodynamic Functions for Solubility of 1-Hydroxybenzotriazole in Sixteen Solvents at Temperatures from (278.15 to 313.15) K and Mixing Property of Mixtures. *Journal of Chemical & Engineering Data* 2017, *62*, 2191-2197.

[27] Jouyban, A., Martinez, F., Acree, W. E., Correct derivation of a combined version of the Jouyban–Acree and van't Hoff model and some comments on 'Determination and correlation of the solubility of myricetin in ethanol and water mixtures from 288.15 to 323.15 K'. *Physics and Chemistry of Liquids* 2017, *55*, 131-140.

[28] Lange, Y., Ory, D. S., Ye, J., Lanier, M. H., *et al.*, Effectors of rapid homeostatic responses of endoplasmic reticulum cholesterol and 3-hydroxy-3-methylglutaryl-CoA reductase. *J Biol Chem* 2008, *283*, 1445-1455.

[29] Smallwood, I. M., *Handbook of organic solvent properties*, Amold, London 1996.

[30] Barton, A. F. M., *Handbook of Solubility Parameters and Other Cohesion Parameters*, CRC Press Boca Raton, F.L 1991

[31] Bosque, R., Sales, J., Polarizabilities of Solvents from the Chemical Composition. *Journal of Chemical Information and Computer Sciences* 2002, *42*, 1154-1163.

[32] Chen, W., Su, B., Xing, H., Yang, Y., Ren, Q., Solubilities of cholesterol and desmosterol in binary solvent mixtures of n-hexane+ethanol. *Fluid Phase Equilibria* 2009, *287*, 1-6.

[33] Liu, X. Y., Sawant, P. D., Micro/Nanoengineering of the Self-Organized Three-Dimensional Fibrous Structure of Functional Materials. *Angewandte Chemie International Edition* 2002, *41*, 3641-3645.

[34] Pan, H. Y., Qiang; Ding, Juping; Yin, Shijie; Tang, Mulin; Wan, Qingyu; Chen, Yan, 2018.

[35] Millard Maienthal , P. J. F., Preparation of Lanosterol From Bromolanosterol. 1955.

[36] Lee, J. W., Thomas, L. C., Schmidt, S. J., Can the Thermodynamic Melting Temperature of Sucrose, Glucose, and Fructose Be Measured Using Rapid-Scanning Differential Scanning Calorimetry (DSC)? *Journal of Agricultural and Food Chemistry* 2011, *59*, 3306-3310.

[37] Kontogeorgis, G. M., Folas, G. K., *Thermodynamic Models for Industrial Applications: From Classical and Advanced Mixing Rules to Association Theories*, Wiley 2009.

[38] Sridhar, S., Moulik, S., *Membrane Processes: Pervaporation, Vapor Permeation and Membrane Distillation for Industrial Scale Separations*, Wiley 2018.

[39] Barzegar-Jalali, M., Rahimpour, E., Acree, W. E., Jouyban, A., A global version of modified Wilson model for solubility prediction of drugs in methanol + water mixtures. *Journal of Molecular Liquids* 2018, *269*, 609-618.

APPENDIX

SUPPORTING INFORMATION OF II. SOLUBILITY OF LANOSTEROL IN ORGANIC SOLVENTS AND IN WATER-ALCOHOL MIXTURES AT 101.8 KPA

Table S1. The solubility parameter, polarity and polarizability of solvents

Solvent	Solubility Parameter[a] $(J \cdot cm^{-3})^{0.5}$	Polarity[a] (Water 100)	Polarizability[b] ($Å^3$)
Methanol	29.52	76.2	3.26
Ethanol	26.58	65.4	5.13
Isopropanol	23.8	54.6	7.14
n-propanol	24.5	61.7	7.23
Acetone	20.05	35.5	6.47
Acetonitrile	24.09	46	4.44
DMF	23.97	40.4	7.93
DMSO	26.33	44.4	8.03
Ethyl Acetate	18.35	23	8.87
n-Hexane	14.9	0.9	11.94

[a] Taken from Reference 1 and 2.

[b] Taken from Reference 3.

Table S2. Parameters for the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models, Root Mean Square Deviation(*RMSD*) and the relative average deviation (*RAD*) for the solubility of lanosterol (x_3) in water-methanol binary mixtures[a]

water(1)-methanol(2)								
T/K	w_2	x_3^{exp}	$x_3^{cal\text{-}apb}$	$x_3^{cal\text{-}van}$		Apelblat–Jouyban-Acree		van't Hoff–Jouyban-Acree
296.67	0.00	3.25E-09	2.27E-09	5.44E-10	A_1	-497.83	A_1	10.81
306.81	0.00	5.30E-09	2.51E-09	1.57E-09	B_1	20730.00	B_1	-9537.74
316.94	0.00	9.33E-09	2.97E-09	4.25E-09	C_1	71.67	A_2	9.36
328.97	0.00	1.75E-08	3.92E-09	1.28E-08	A_2	-823.63	B_2	-5620.29
338.78	0.00	2.81E-08	5.19E-09	2.96E-08	B_2	32600.00	J_0	-5277.64
296.67	0.040	3.53E-09	3.55E-09	1.14E-09	C_2	123.70	J_1	126.75
306.81	0.040	5.95E-09	3.99E-09	3.22E-09	J_0	-4922.56	J_2	8573.20
316.94	0.040	1.33E-08	4.80E-09	8.47E-09	J_1	-685.88		
328.97	0.040	3.02E-08	6.48E-09	2.48E-08	J_2	6864.09		
338.78	0.040	6.56E-08	8.74E-09	5.62E-08				
296.67	0.16	3.66E-09	4.37E-09	2.04E-09				
306.81	0.16	6.34E-09	5.34E-09	5.62E-09				
316.94	0.16	1.51E-08	7.01E-09	1.45E-08				
328.97	0.16	3.70E-08	1.05E-08	4.12E-08				
338.78	0.16	7.33E-08	1.54E-08	9.16E-08				
296.67	0.44	3.83E-09	3.69E-09	1.34E-09				
306.81	0.44	7.91E-09	5.62E-09	3.68E-09				
316.94	0.44	2.06E-08	9.21E-09	9.48E-09				

Table S2. Parameters for the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models, Root Mean Square Deviation(*RMSD*) and the relative average deviation (*RAD*) for the solubility of lanosterol (x_3) in water-methanol binary mixtures (cont.)

water(1)-methanol(2)

T/K	w_2	x_3^{exp}	$x_3^{\text{cal-apb}}$	$x_3^{\text{cal-van}}$	Apelblat–Jouyban-Acree	van't Hoff–Jouyban-Acree
328.97	0.44	4.72E-08	1.80E-08	2.70E-08		
338.78	0.44	1.01E-07	3.27E-08	6.01E-08		
296.67	0.65	6.49E-08	8.24E-08	3.46E-08		
306.81	0.65	1.01E-07	1.33E-07	8.45E-08		
316.94	0.65	1.98E-07	2.33E-07	1.95E-07		
328.97	0.65	4.74E-07	4.97E-07	4.92E-07		
338.78	0.65	1.00E-06	9.79E-07	9.98E-07		
296.67	0.88	1.30E-05	1.77E-05	1.45E-05		
306.81	0.88	2.94E-05	2.86E-05	2.87E-05		
316.94	0.88	5.47E-05	5.08E-05	5.43E-05		
328.97	0.88	0.00011	0.000111	0.00011		
277.78	1.00	2.70E-05	3.69E-05	1.9E-05		
288.15	1.00	4.86E-05	5.03E-05	3.93E-05		
298.15	1.00	8.83E-05	7.7E-05	7.56E-05		
303.05	1.00	0.00011	9.87E-05	0.000103		
307.55	1.00	0.00013	0.00013	0.00013		
312.15	1.00	0.00016	0.00017	0.00018		
316.95	1.00	0.00021	0.00023	0.00023		
321.65	1.00	0.00032	0.00031	0.00030		

Table S2. Parameters for the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models, Root Mean Square Deviation(*RMSD*) and the relative average deviation (*RAD*) for the solubility of lanosterol (x_3) in water-methanol binary mixtures (cont.)

water(1)-methanol(2)						
T/K	w_2	x_3^{exp}	$x_3^{\text{cal-apb}}$	$x_3^{\text{cal-van}}$	Apelblat–Jouyban-Acree	van't Hoff–Jouyban-Acree
Apelblat–Jouyban-Acree	*RAD*	0.39				
	RMSD	1.88E-06				
van't Hoff–Jouyban-Acree	*RAD*	0.21				
	RMSD	2.59E-06				

[a] w_2 is mass fraction of alcohols in a water-alcohol binary system free of lanosterol; $x_3^{\text{cal-apb}}$ and $x_3^{\text{cal-van}}$ are calculated lanosterol mole fractions by the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models respectively.

Table S3. Parameters for the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models, Root Mean Square Deviation(*RMSD*) and the relative average deviation (*RAD*) for the solubility of lanosterol (x_3) in water-ethanol binary mixtures.[a]

water(1)-ethanol(2)								
T/K	w_2	x_3^{exp}	$x_3^{\text{cal-apb}}$	$x_3^{\text{cal-van}}$		Apelblat–Jouyban-Acree		van't Hoff–Jouyban-Acree
278.15	0.00	1.43E-09	4.37E-09	9E-09	A_1	-507.613	A_1	-9.35289
288.15	0.00	1.94E-09	6.33E-09	1.22E-08	B_1	19136.32	B_1	-2721.48
298.15	0.00	4.17E-09	9.68E-09	1.62E-08	C_1	74.45794	A_2	2.542052
308.15	0.00	6.21E-09	1.70E-08	2.22E-08	A_2	-727.768	B_2	-2785.93

Table S3. Parameters for the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models, Root Mean Square Deviation(*RMSD*) and the relative average deviation (*RAD*) for the solubility of lanosterol (x_3) in water-ethanol binary mixtures (cont.)

water(1)-ethanol(2)

T/K	w_2	x_3^{exp}	$x_3^{cal\text{-}apb}$	$x_3^{cal\text{-}van}$		Apelblat–Jouyban-Acree		van't Hoff–Jouyban-Acree
318.15	0.00	1.25E-08	2.82E-08	2.81E-08	B_2	30018.84	J_0	-1969.19
278.15	0.040	1.73E-08	1.28E-08	1.19E-08	C_2	108.8534	J_1	-14.3242
288.15	0.040	2.79E-08	1.56E-08	1.65E-08	J_0	-2745.25	J_2	5998.03
298.15	0.040	4.57E-08	2.05E-08	2.24E-08	J_1	1714.521		
308.15	0.040	7.64E-08	2.88E-08	2.97E-08	J_2	10612.25		
318.15	0.040	1.06E-07	4.27E-08	3.88E-08				
278.15	0.25	6.07E-08	1.01E-07	6.54E-08				
288.15	0.25	8.96E-08	1.22E-07	9.31E-08				
298.15	0.25	1.35E-07	1.61E-07	1.29E-07				
308.15	0.25	1.80E-07	2.27E-07	1.76E-07				
318.15	0.25	2.25E-07	3.42E-07	2.35E-07				
278.15	0.54	1.73E-07	1.93E-07	4.74E-07				
288.15	0.54	2.87E-07	2.48E-07	7.10E-07				
298.15	0.54	4.42E-07	3.50E-07	1.03E-06				
308.15	0.54	8.38E-07	5.34E-07	1.47E-06				
318.15	0.54	1.20E-06	8.72E-07	2.05E-06				
278.15	0.88	0.00022	0.00034	0.00025				
288.15	0.88	0.00033	0.00038	0.00034				
298.15	0.88	0.00047	0.00046	0.00047				
308.15	0.88	0.00069	0.00063	0.00062				
318.15	0.88	0.00077	0.00093	0.00081				

Table S3. Parameters for the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models, Root Mean Square Deviation(*RMSD*) and the relative average deviation (*RAD*) for the solubility of lanosterol (x_3) in water-ethanol binary mixtures (cont.)

water(1)-ethanol(2)						
T/K	w_2	x_3^{exp}	$x_3^{\text{cal-apb}}$	$x_3^{\text{cal-van}}$	Apelblat–Jouyban-Acree	van't Hoff–Jouyban-Acree
318.05	1.00	0.0021	0.0021	0.0020		
312.35	1.00	0.0017	0.0017	0.0017		
307.15	1.00	0.0014	0.0014	0.0015		
302.85	1.00	0.0012	0.0012	0.0013		
298.25	1.00	0.0010	0.0010	0.0011		
293.45	1.00	0.00092	0.00091	0.0010		
285.74	1.00	0.00082	0.00080	0.00074		
277.95	1.00	0.00073	0.00075	0.00056		
Apelblat–Jouyban-Acree	*RAD*	0.32				
	RMSD	2.88E-05				
van't Hoff–Jouyban-Acree	*RAD*	0.38				
	RMSD	3.98E-05				

[a] w_2 is mass fraction of alcohols in a water-alcohol binary system free of lanosterol; $x_3^{\text{cal-apb}}$ and $x_3^{\text{cal-van}}$ are calculated lanosterol mole fractions by the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models respectively.

Table S4. Parameters for the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models, Root Mean Square Deviation(*RMSD*) and the relative average deviation (*RAD*) for the solubility of lanosterol (x_3) in water-isopropanol binary mixtures

water(1)-isopropanol(2)								
T/K	w_2	x_3^{exp}	$x_3^{cal\text{-}apb}$	$x_3^{cal\text{-}van}$		Apelblat–Jouyban-Acree		van't Hoff–Jouyban-Acree
278.15	0.00	1.43E-09	4.37E-09	8.43E-13	A_1	-507.61	A_1	5.759314
288.15	0.00	1.94E-09	6.33E-09	8.75E-13	B_1	19136.32	B_1	-3771.12
298.15	0.00	4.17E-09	9.68E-09	9.06E-13	C_1	74.45	A_2	-26.6873
308.15	0.00	6.21E-09	1.70E-08	9.41E-13	A_2	16.99	B_2	-330.517
318.15	0.00	1.25E-08	2.82E-08	9.69E-13	B_2	-4281.89	J_0	7328.281
278.15	0.040	8.79E-09	3.68E-09	1.55E-11	C_2	-1.67	J_1	-6306.84
288.15	0.040	1.23E-08	4.75E-09	1.53E-11	J_0	1632.95	J_2	3343.804
298.15	0.040	1.41E-08	6.56E-09	1.5E-11	J_1	-1464.48		
308.15	0.040	1.69E-08	9.59E-09	1.48E-11	J_2	-1554.38		
318.15	0.040	2.21E-08	1.47E-08	1.46E-11				
278.15	0.25	7.97E-08	7.64E-08	2.57E-07				
288.15	0.25	9.02E-08	1.02E-07	2.29E-07				
298.15	0.25	1.07E-07	1.42E-07	2.06E-07				
308.15	0.25	1.19E-07	2.06E-07	1.86E-07				
318.15	0.25	1.31E-07	3.10E-07	1.69E-07				
278.15	0.54	1.07E-05	8.03E-06	1.82E-05				
288.15	0.54	1.41E-05	1.10E-05	1.94E-05				

Table S4. Parameters for the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models, Root Mean Square Deviation(*RMSD*) and the relative average deviation (*RAD*) for the solubility of lanosterol (x_3) in water-isopropanol binary mixtures (cont.)

water(1)-isopropanol(2)						
T/K	w_2	x_3^{exp}	$x_3^{cal\text{-}apb}$	$x_3^{cal\text{-}van}$	Apelblat–Jouyban-Acree	van't Hoff–Jouyban-Acree
298.15	0.54	2.08E-05	1.53E-05	2.05E-05		
308.15	0.54	2.51E-05	2.17E-05	2.17E-05		
318.15	0.54	3.10E-05	3.10E-05	2.28E-05		
278.15	0.88	0.00016	0.00019	0.00020		
288.15	0.88	0.00030	0.00029	0.00028		
298.15	0.88	0.00041	0.00044	0.00039		
308.15	0.88	0.00052	0.00064	0.00053		
318.15	0.88	0.00071	0.00093	0.00071		
278.07	1.00	0.00047	0.00041	0.00041		
287.81	1.00	0.00060	0.00065	0.0006		
298.84	1.00	0.00095	0.00105	0.0010		
307.27	1.00	0.0016	0.0015	0.0015		
318.93	1.00	0.0023	0.0023	0.0023		
Apelblat – Jouyban-Acree	*RAD*	0.36				
	RMSD	0.00037				
van't Hoff– Jouyban-Acree	*RAD*	0.23				
	RMSD	6.51E-06				

[a] w_2 is mass fraction of alcohols in a water-alcohol binary system free of lanosterol; $x_3^{cal\text{-}apb}$ and $x_3^{cal\text{-}van}$ are calculated lanosterol mole fractions by the Apelblat–Jouyban-Acree and van't Hoff–Jouyban-Acree models respectively.

Table S5. Interaction Parameters for the NRTL and UNIQUAC Models

Solvent	NRTL		
	Δg_{12}	Δg_{21}	a
Methanol	3786.04	-1893.02	0.20
Ethanol	5866.49	-2933.24	0.20
Isopropanol	5929.32	-2964.66	0.20
n-Propanol	5389.93	-2694.96	0.20
Acetone	4027.24	-2013.62	0.20
Acetonitrile	6080.32	-3040.16	0.20
Ethyl Acetate	3866.78	-1933.39	0.20
Solvent	UNIQUAC		
	Δu_{12}	Δu_{21}	
Methanol	162.67	-81.34	
Ethanol	90.00	-45.00	
Isopropanol	153.36	-76.68	
n-Propanol	147.01	-73.51	
Acetone	83.04	-41.53	
Acetonitrile	136.14	-68.07	
Ethyl Acetate	62.43	-31.22	

Table S6. The Calculated Values for ΔG_{mix} Gibbs Free Energy of Mixing, ΔH_{mix} Enthalpy of Mixing and ΔS_{mix} Entropy of Mixing

T/K	ΔG_{mix}	ΔH_{mix}	ΔS_{mix}
	J·mol^{-1}	J·mol^{-1}	J·mol^{-1}
Methanol			
277.78	-1.19	0.00040	0.0043
288.15	-2.16	0.00072	0.0070
298.15	-3.93	0.0013	0.013
303.05	-4.82	0.0016	0.016
307.55	-5.70	0.0019	0.019
312.15	-7.11	0.0024	0.023
316.95	-9.54	0.0032	0.030
321.65	-14.28	0.0048	0.044
Ethanol			
277.95	-14.16	0.013	0.051
285.74	-21.25	0.020	0.074
293.45	-30.10	0.028	0.10
298.25	-37.17	0.035	0.12
302.85	-47.83	0.045	0.15
307.15	-54.82	0.051	0.17
312.35	-68.29	0.064	0.21
318.05	-87.13	0.082	0.27
Isopropanol			
278.07	-18.45	0.014	0.066
287.81	-23.84	0.018	0.083
298.84	-38.05	0.029	0.13
307.27	-63.57	0.048	0.21

Table S6. The Calculated Values for ΔG_{mix} Gibbs Free Energy of Mixing, ΔH_{mix} Enthalpy of Mixing and ΔS_{mix} Entropy of Mixing (cont.)

T/K	ΔG_{mix}	ΔH_{mix}	ΔS_{mix}
	J·mol^{-1}	J·mol^{-1}	J·mol^{-1}
318.93	-92.89	0.069	0.29
n-propanol			
278.07	-32.16	0.033	0.12
287.81	-51.60	0.052	0.18
298.84	-75.64	0.076	0.25
307.27	-110.31	0.11	0.36
318.93	-174.39	0.17	0.55
Acetone			
278.65	-19.19	0.015	0.070
283.75	-25.20	0.020	0.090
288.45	-30.50	0.024	0.11
293.25	-36.92	0.029	0.13
298.05	-45.56	0.036	0.15
302.75	-55.89	0.044	0.18
307.55	-68.47	0.054	0.22
318.36	-103.74	0.082	0.33
Acetonitrile			
277.09	-1.38	-0.00021	0.0050
288.55	-2.80	-0.00041	0.010
293.35	-4.10	-0.00060	0.014
298.15	-5.45	-0.00080	0.018
303.05	-6.75	-0.0010	0.022

Table S6. The Calculated Values for ΔG_{mix} Gibbs Free Energy of Mixing, ΔH_{mix} Enthalpy of Mixing and ΔS_{mix} Entropy of Mixing (cont.)

T/K	ΔG_{mix}	ΔH_{mix}	ΔS_{mix}
	J·mol^{-1}	J·mol^{-1}	J·mol^{-1}
307.55	-8.48	-0.0012	0.028
312.25	-11.33	-0.0016	0.036
317.05	-16.49	-0.0024	0.052
Ethyl Acetate			
277.67	-21.42	-0.0060	0.077
288.55	-35.80	-0.010	0.12
293.25	-46.22	-0.013	0.16
298.05	-52.65	-0.015	0.18
302.85	-65.88	-0.019	0.22
307.45	-79.45	-0.022	0.26
318.64	-123.16	-0.035	0.39

Table S7. Solubility of lanosterol (3) in water (1)-methanol (2) binary mixture predicted by the Modified Wilson compare with experimental data

	296.67 K			306.81 K			316.94 K		
w_2	$x^{exp}{}_3$	$x^{pre}{}_3$	*RD %*	$x^{exp}{}_3$	$x^{pre}{}_3$	*RD%*	$x^{exp}{}_3$	$x^{pre}{}_3$	*RD %*
0.04	3.53E-09	1.92E-09	45.55	5.95E-09	3.69E-09	37.89	1.33E-08	6.34E-09	52.17
0.16	3.66E-09	1.58E-09	56.76	6.34E-09	3.28E-09	48.17	1.51E-08	6.34E-09	0.58
0.44	3.83E-09	2.82E-08	636.15	7.91E-09	5.84E-08	638.05	2.06E-08	1.21E-07	4.90
0.65	6.49E-08	4.95E-07	662.08	1.01E-07	9.62E-07	851.77	1.98E-07	1.95E-06	8.88
0.88	1.30E-05	1.30E-05	0.34	2.94E-05	2.32E-05	21.07	5.47E-05	4.48E-05	0.18

	328.97 K			338.78 K		
w_2	$x^{exp}{}_3$	$x^{pre}{}_3$	*RD%*	$x^{exp}{}_3$	$x^{pre}{}_3$	*RD%*
0.04	3.02E-08	1.33E-08	56.08	6.56E-08	2.36E-08	64.03
0.16	3.70E-08	1.55E-08	57.98	7.33E-08	3.22E-08	56.11
0.44	4.72E-08	3.30E-07	599.74	1.01E-07	7.73E-07	665.95
0.65	4.74E-07	5.13E-06	982.59	1.00E-06	1.19E-05	1084.95
0.88	0.00011	0.00011	1.16			

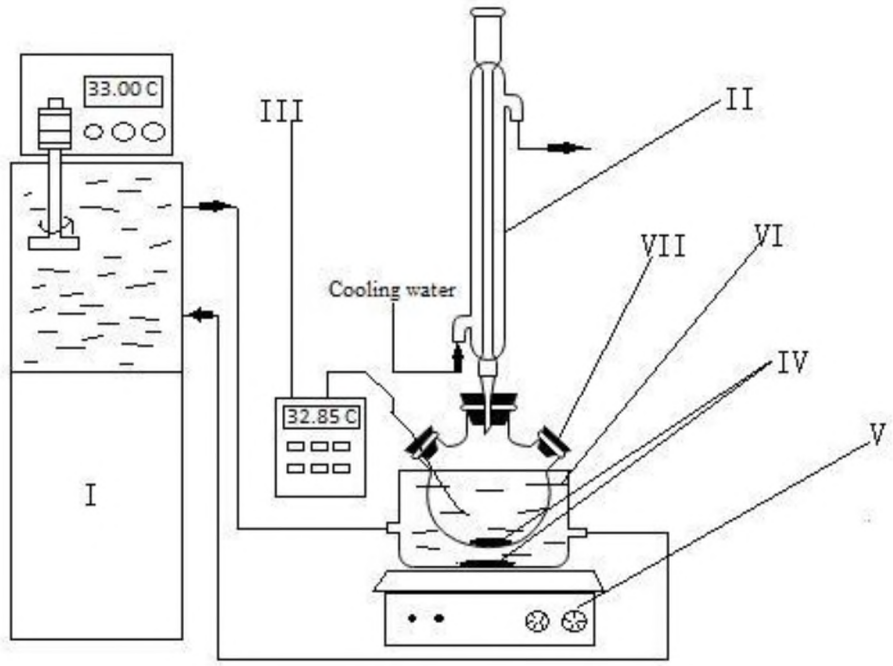

Figure S1. Schematic diagram of the experimental set up: I, water circulating bath; II, condenser; III, thermocouple; IV, magnetic stirrers; V, magnetic stirrer controller; VI, water jacketed three necks glass vessel; VII, sampling port

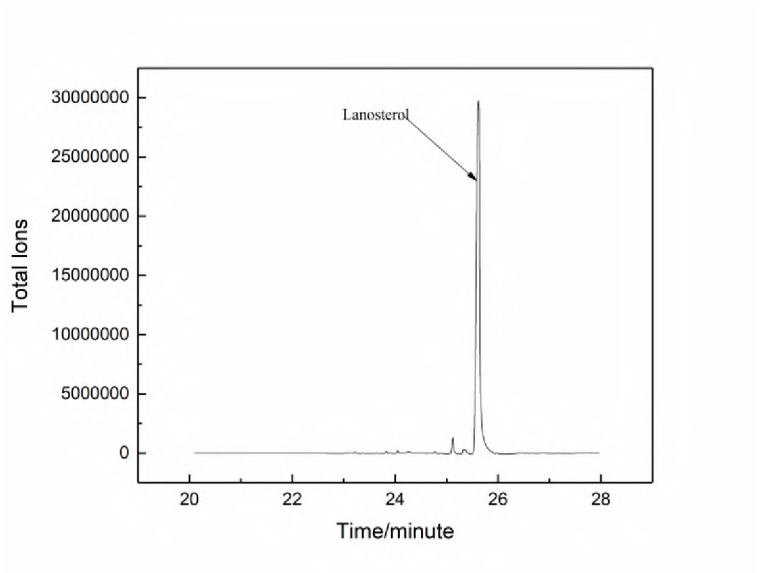

Figure S2. Total ions chromatogram of purified lanosterol in methanol by GC-MS. The retention time of lanosterol is 25.62 minutes and the purity of lanosterol is >95%

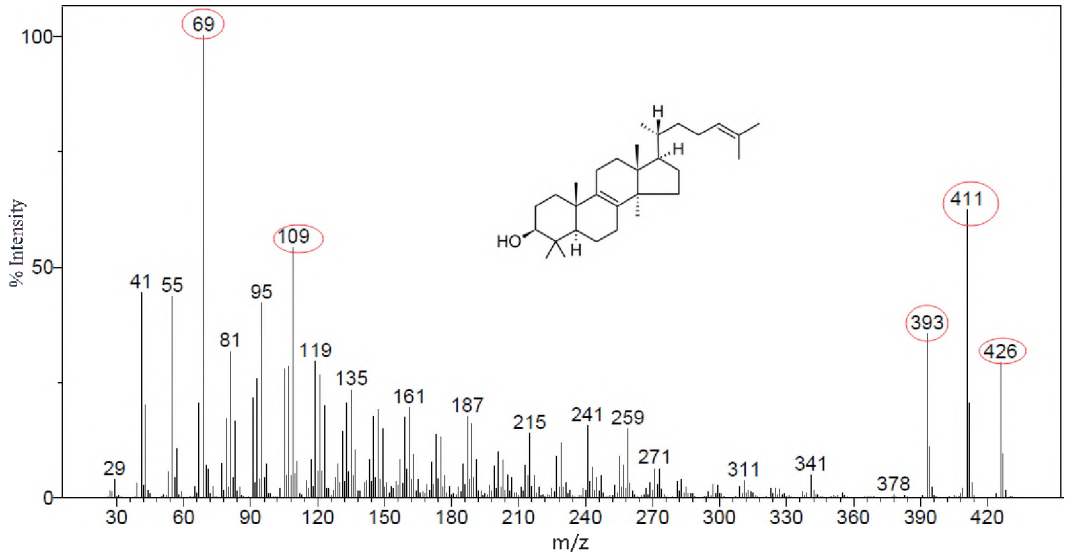

Figure S3. MS scan of lanosterol in methanol at 25.62 minutes. Ion groups (m/z) are found as 69,109,393,411 and 426 which indicate by a red circle, respectively. MS scan results were compared with data from reference 4 to confirm that the sample is lanosterol

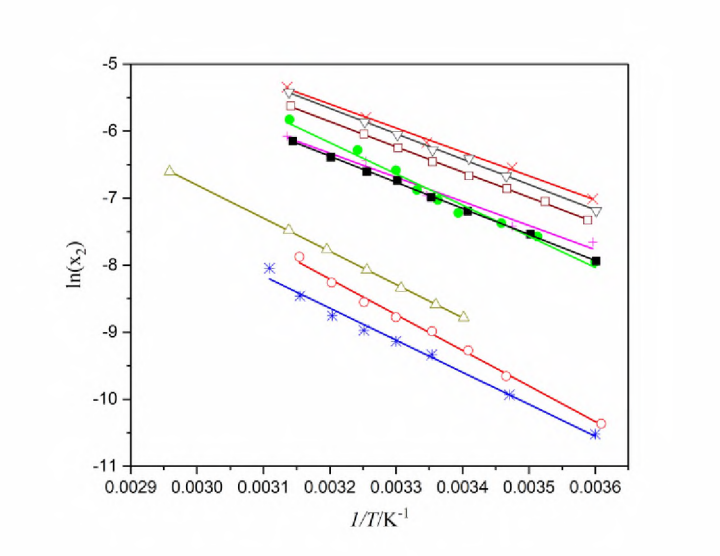

Figure S4. van't Hoff plots of ln(x_2) versus $1/T$ in different solvent. ▼,methanol; ○,acetonitrile;△, DMSO; ●, DMF; ■, ethanol; □, acetone; ▽, ethyl acetate; ▲, isopropanol; and ×, *n*-propanol. The solid lines are the solubilities fitted by the modified Apelblat equation

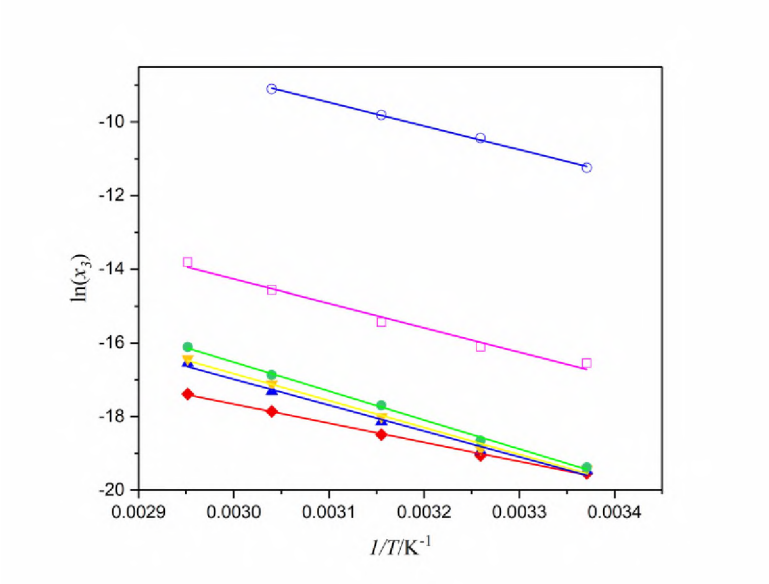

Figure S5. van't Hoff plots of ln(x_3) versus $1/T$ in water (1)-methanol (2) mixtures. ◆, water; ▲, v_2=0.050; ▼, v_2=0.20; ●, v_2=0.50; □, v_2=0.70;and ○, v_2=0.90. The solid lines are the solubilities fitted by the modified Apelblat equation

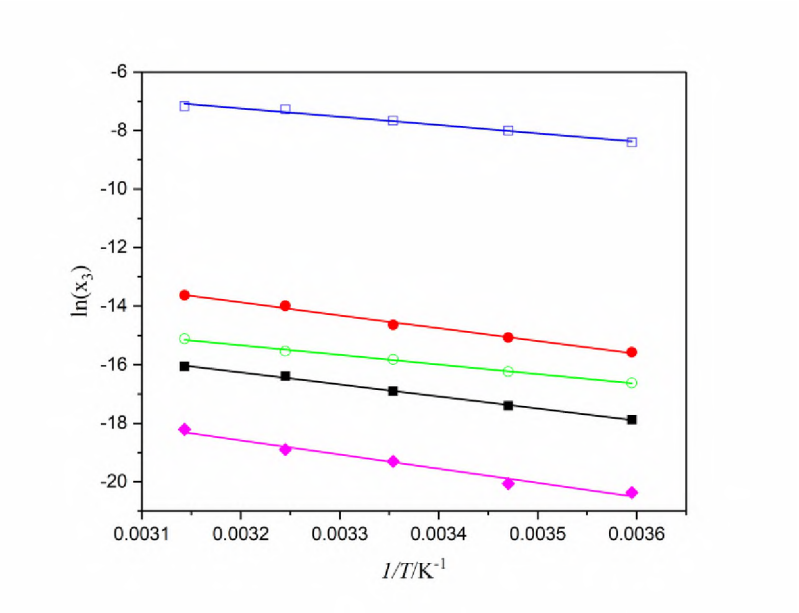

Figure S6. van't Hoff plots of ln(x_3) versus $1/T$ in water (1)-ethanol (2) mixtures. ◆, water; ■, v_2=0.050; ○, v_2=0.30; ●, v_2=0.60; □, v_2=0.90. The solid lines are the solubilities fitted by the modified Apelblat equation

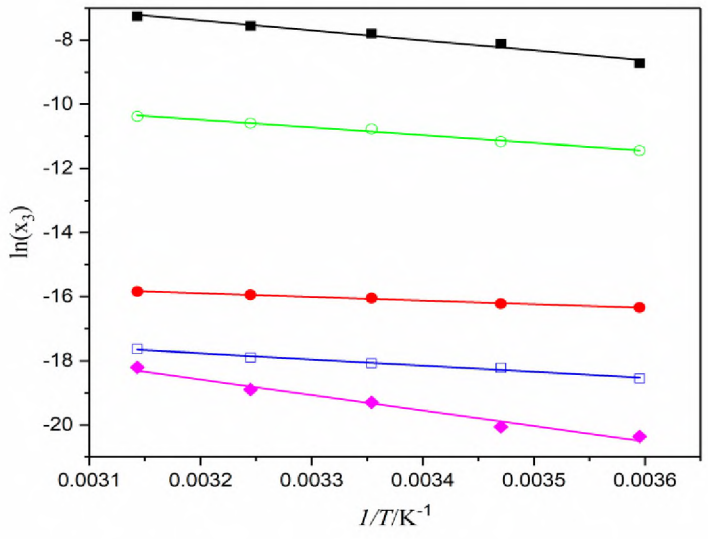

Figure S7. van't Hoff plots of ln(x_3) versus $1/T$ in water (1)-isopropanol (2) mixtures. ◆, water; □ v_2=0.050; ●, v_2=0.30; ○, v_2=0.60;■, v_2=0.90. The solid lines are the solubilities fitted by the modified Apelblat equation

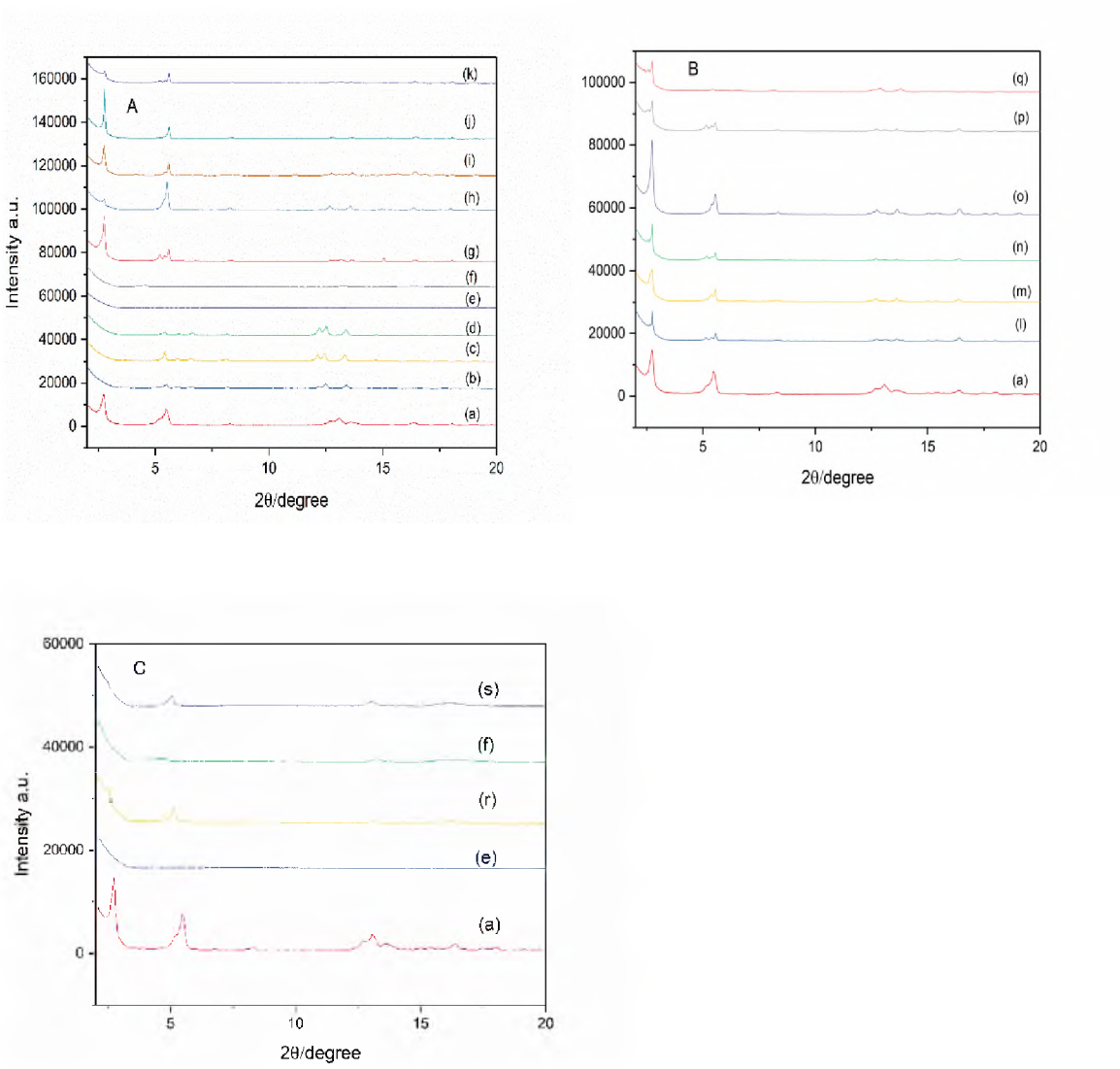

Figure S8. X-ray diffraction patterns from 2° to 20° 2θ/ degree of lanosterol powder and lanosterol precipitates at equilibrium with lanosterol solutions in pure solvents and water-alchol mixtures. A) (a) powder lanosterol before solubility experiments; (b) ethanol; (c) isopropanol; (d) *n*-propanol; (e) DMSO; (f) DMF; (g) actonitrile; (h) methanol; (i) acetone; (j) ethyl acetate; (k) water; B) (l) 5%(v/v) ethanol; (m) 90% ethanol; (n) 5% methanol; (o) 90% methanol; (p) 5% isopropanol; (q) 90% isopropanol; C) lanosterol precipitates at equilibrium with solutions of lanosterol in DMSO and DMF vacuum-oven dried 24 hours; (a) lanosterol powder before solubility experiments; (e) wet lanosterol at equilibrium with a lanosterol solution in DMSO; (r) dried lanosterol at equilibrium with a lanosterol solution in DMSO; (s), dried lanosterol at equilibrium with a lanosterol solution in DMF

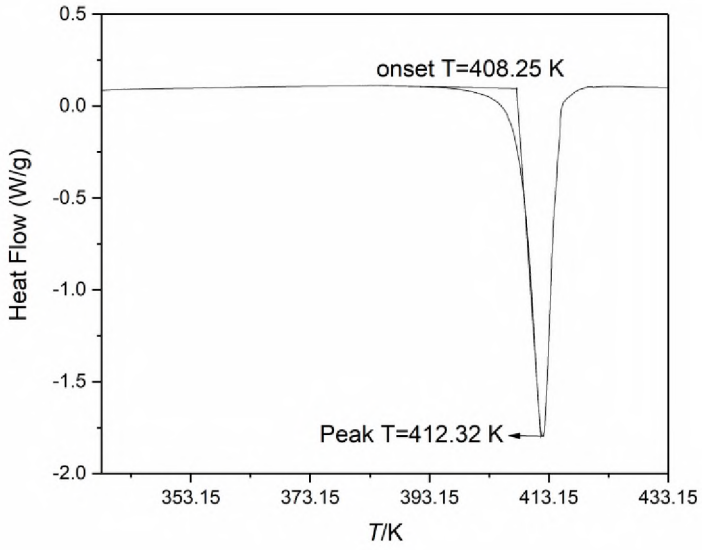

Figure S9. Differential scanning calorimetry (DSC) of lanosterol from 338.15 K-433.15 K. The onset T_m is 408.27 K with an uncertainty $u(T_m)$ = 0.50 K; the peak of T_m is 412.39 K with an uncertainty $u(T_m)$ = 0.60 K; the enthalpy of fusion ΔH_{fus} of lanosterol is 23.61 kJ with an uncertainty u (ΔH_{fus}) = 0.13kJ·mol^{-1}. DSC experiments were done in triplicate

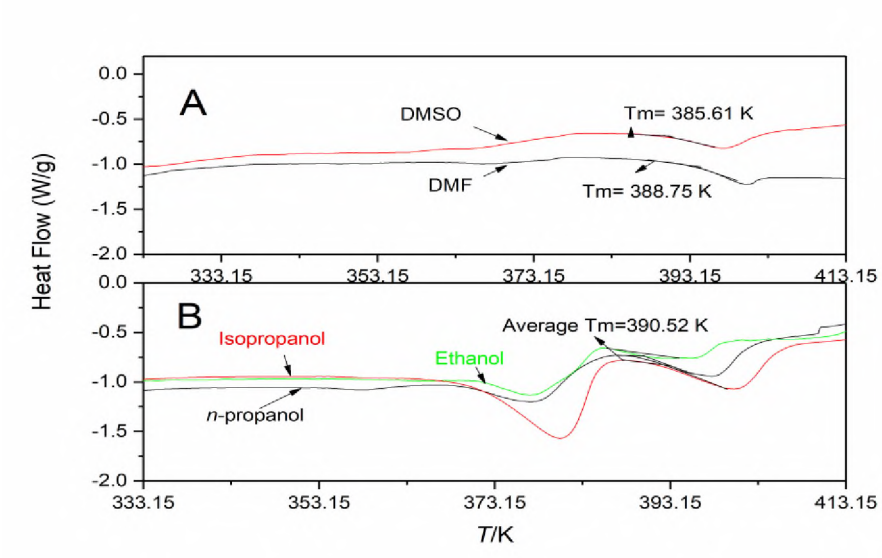

Figure S10. Differential scanning calorimetry (DSC) of lanosterol after solubility measurements. A), dried precipitate at equilibrium with a lanosterol solution in DMSO and DMF; B), dried precipitate at equilibrium with a lanosterol solution in ethanol, isopropanol and *n*-propanol

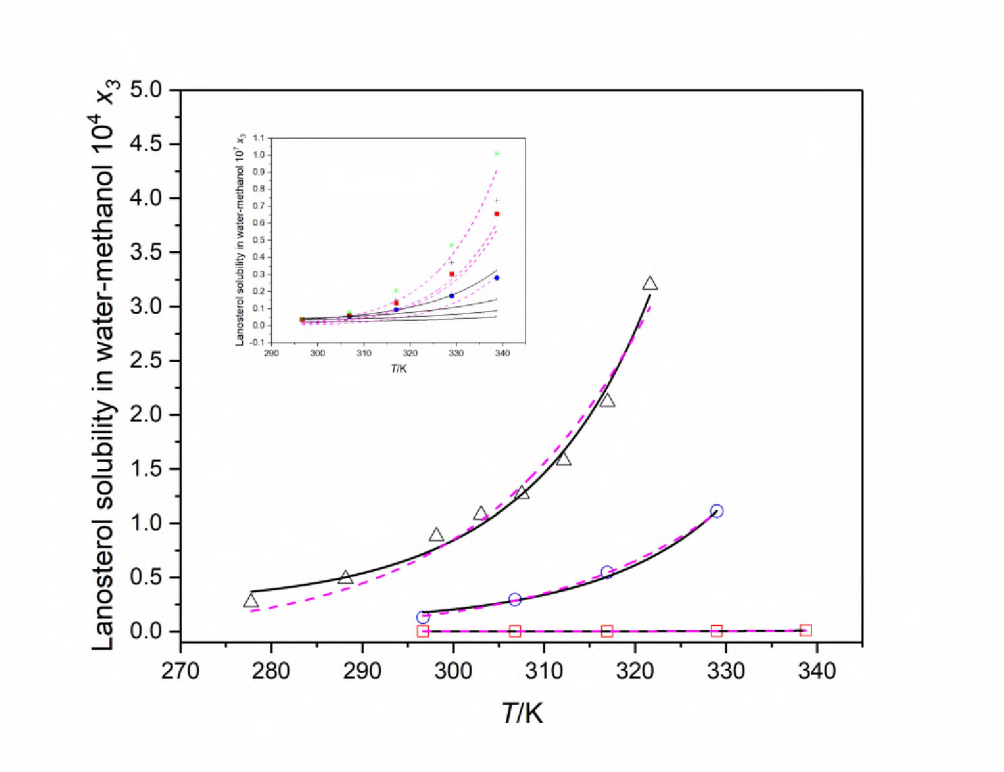

Figure S11. The mole fraction of lanosterol (3), x_3, in water and water (1)-methanol (2) mixture as a function of temperature. □, v_2=0.70; ○, v_2=0.90; △, v_2=1.00. The insert shows x_3 in ● , water v_2=0.00; ■, v_2=0.050; ▲, v_2=0.20; ▼, v_2=0.50; The solid lines are the solubilities fitted by the Apelblat–Jouyban-Acree model and the dash lines are the solubilities fitted by van't Hoff–Jouyban-Acree model

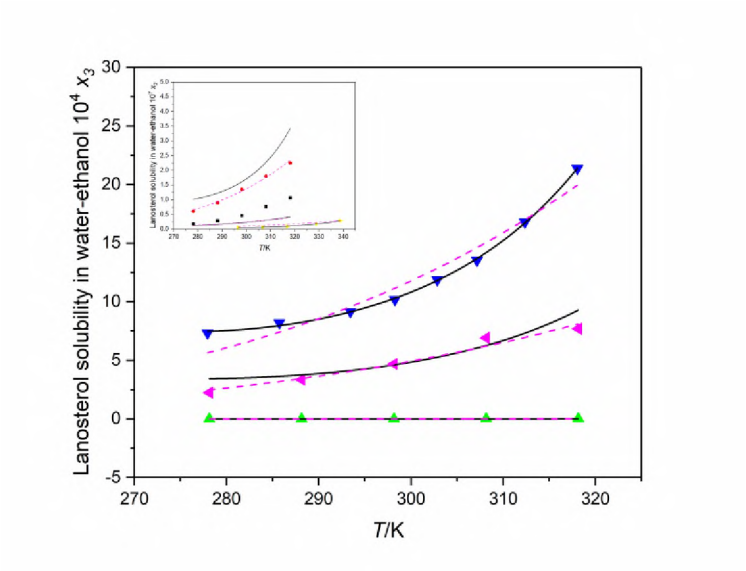

Figure S12. The mole fraction of lanosterol (3) x_3 in water and water (1)-ethanol (2) mixture as a function of temperature. ▲, v_2=0.60; ◀, v_2=0.90; ▼, v_2= 1.00. The insert shows x_3 in :▶, water; ■, v_2=0.050; and ●, v_2=0.30. The solid lines are the solubilities fitted by the Apelblat–Jouyban-Acree model and the dash lines are the solubilities fitted by van't Hoff–Jouyban-Acree model

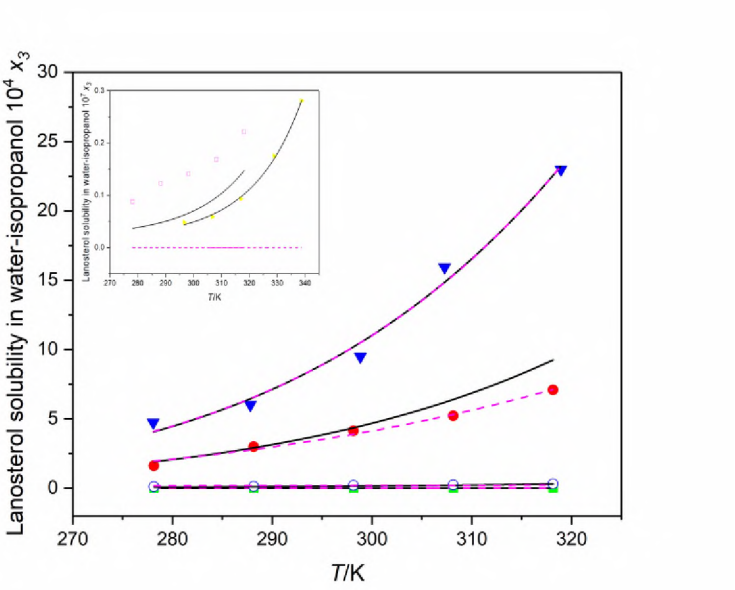

Figure S13. The mole fraction of lanosterol (3) x_3 in water and water (1)-isopropanol (2) mixtures as a function of temperature. ■, v_2=0.30; ○, v_2=0.60; ●, v_2=0.90; and ▼, v_2= 1.00. The insert shows x_3 in ▶, water; and □, v_2=0.050. The solid lines are the solubilities fitted by the Apelblat–Jouyban-Acree model and the dash lines are the solubilities fitted by van't Hoff–Jouyban-Acree model

REFERENCES

[1] Smallwood, I. M. *Handbook of Organic Solvent Properties*. Amold: London, 1996.

[2] Barton, A. F. M. *Handbook of Solubility Parameters and Other Cohesion Parameters*. CRC Press: Boca Raton, F.L, 1991

[3] Bosque, R., Sales, J. Polarizabilities of Solvents from the Chemical Composition. *J. Chem. Inf. Comput. Sci.* 2002, *42*, 1154-1163.

[4] Lange, Y., Ory, D. S., Ye, J., Lanier, M. H., Hsu, F. F., Steck, T. L. Effectors of Rapid Homeostatic Responses of Endoplasmic Reticulum Cholesterol and 3-hydroxy-3-methylglutaryl-CoA Reductase. *J. Biol. Chem.* 2008, *283*, 1445-55

III. IS IT POSSIBLE TO PREVENT AGGREGATION OF ALPHA CRYSTALLIN BY ADDING STEROLS?

ABSTRACT

Cataract is the main cause of blindness in the world. By age 75, half of the Americans will have cataracts according to the National Eye Institute. The only available treatment consists of replacing the damaged lenses by artificial ones. Although the surgery is safe and it corrects vision problems, not every patient has easy to access to the surgery especially in developing countries. In 2015 two reports were published claiming that lanosterol (LAN) and 25-hydroxycholesterol (25HD) restored lens' clarity. Since then, there is almost an equal number of publications reporting restoration of lenses clarity by lanosterol and 25-hydroxycholesterol as those refuting those findings. There is a broader question to be asked: are sterols able to prevent protein aggregation in general? Previous studies were all focused on restoring the lenses' transparency but the effects of the two sterols on the α-crystallin aggregation process has not been investigated. In this study we showed these two sterols fail to prevent α-crystallin heat induced aggregation. Furthermore, sterols at high concentration actually promote α-crystallin aggregation but the α-crystallin chaperone activity seems to remain intact. FTIR and CD spectra show that the secondary and tertiary structures of α-crystallin does not change significantly in the present of the two sterols. Cu^{2+} binding experiments and bis-ANS hydrophobic fluorescent assay further indicate that there is no interactions between α-crystallin and those two sterols. Our results show no evidence to support specific interactions between

α-crystallin and those two sterols and that the lost in chaperone activity is caused by aggregation of α-crystallin.

1. INTRODUCTION

Cataracts are the main cause of blindness in the world. By age 75, half of the Americans have cataracts according to the National Eye Institute.[1] The only available treatment presented by previous literature was replacing the damaged lenses by artificial ones. Although the surgery is safe and it corrects vision problems, not every patient, especially in developing countries, has easy to access to the procedure.[2]

The lens has an onion-like layered structures. Outward facing lens edges have mono-layer of epithelial cells that differentiate to new fiber cells during lens development over the lifetime of a individual[2]. To maintain lens transparency, fiber cells lack of blood vessels and the sub-cellular structures of the fiber cells were removed during differentiation[3]. Lenses' high refractive indexes are caused by the high concentration of crystallins expressed in fiber cells. Furthermore, only the epithelial cells have metabolic activity; therefore, lens fiber cells cannot participate in protein turnover and repair.[4] Among lens crystallins, α-crystallin serves as protein chaperone that prevents aggregation of other crystallins. The chaperon activity of α-crystallin decreases with age, so the lenses lose protection from degradation and oxidation of lens proteins with a consequent increase in the scattering of light, thus forming cataracts.[5, 6]

α-crystallin is a hetero-dimer made by αA and αB crystallins that is usually present as an oligomer of 10-15 hetero-dimers depending on the conditions.[7, 8] α-

crystallin dimer are connected by two pairs of salt bridges at residues 120R and 109D. The alpha crystallin domain (ACD), which also involves in the formation of dimer, consists of a β-sandwich structure of 6 β-strands. αA and αB crystallin belong to the small heat shock protein (sHsp) super family[9]. Members of the sHsp are found in all forms of life and have a highly conserved alpha crystallin domain structure across species. The sHsps are molecular chaperones that bind to unfolded or partially unfolded proteins, preventing their interaction with other unstable proteins. The mini αA crystallin is a recombinant peptide of αA crystallin 70-88 segment (mini-αA 70–88 KFVIFLDVKHFSPEDLTVK), and it shows chaperone activity.[10] The sHsp binds to unstable (partly denatured) proteins and prevents them from aggregating, but the denatured protein can only be restored to its native state in cooperation with other heat shock proteins, such as ATP-driven Hsp70.[11]

In 2015, Zhao *et al.* [12]and Makley *et al.*[13] reported that lanosterol (LAN) and 25-hydroxycholesterol (25HD) restored the lens clarity. They claim those two sterols interact with α-crystallin and enhanced the α-crystallin chaperone activity to such an extent that the chaperone dissolves or disaggregates lens protein aggregates.

Following the studies of lanosterol and 25-hydroxycholesterol to restore lens' transparency, other researchers tried to repeat those experiments or collect additional experimental evidence about the activity of those sterols. The results are mixed. Chen *et al.*[14] showed the lanosterol and 25-hydroxycholesterol in 5% DMSO effectively redissolved human cataractous samples *ex vivo* with a EC50 (Half maximal effective concentration) at 10 μM level. They concluded that those two sterols interact with crystallins by different mechanisms, aggregates of all members of crystallins could be

dissolved by lanosterol, but 25-hydroxycholesterol was only specific to α-crystallin, which contradicts Makley *et al.*'s findings[13]. Shen *et.al.*[15] found that, in vitro, lanosterol (40 μM in M199 medium) delayed the occurrence of lens opacity in a lanosterol synthase inhibited rat lens. Xu *et.al.*[16] used 20 μM lanosterol in 1% DMSO to successfully reverse W151R mutant human βB2-crystallin aggregates. Kang *et al.*[17] used all atom molecular dynamics simulation and free energy perturbation techniques to show that lanosterol can bind to the hydrophobic interface of dimers of human γD-crystallins preventing aggregation. Yang *et al.*[18] successfully synthesized a series of lanosterol derivatives and reported that a few of them reversed mutant crystallins induced protein aggregation. Zhou *et al.*[19] reported that lanosterol (200 and 500 μM in PBS) disrupts the fibrillation of amyloid-β peptides besides redissolving crystallin aggregates. They further investigated the interaction of lanosterol and amyloid-β peptides by molecular dynamics simulations, and then they concluded that lanosterol entangles with the core segment of amyloid-β peptides and forms a hydrophobic core through aromatic side chains. Chemerovski-Glikman *et al.*[20] found that 1 mM 25-hydroxycholesterol in 10% DMSO PBS but not lanosterol resulted in ~20% reduction of cataract solution (a homogenized and resuspended crystallin precipitated) turbidity. Despite the encouraging results obtained in *in vitro* experiments, the clinical trails and the *in vivo* experiments using these two sterols have been unsatisfactory. Felici *et.al.*[21] reported that using 5 mM lanosterol in an olive oil eye drop given to patients with idiopathic unilateral juvenile nuclear cataracts failed to dissolve cataract or halt the progress of lens opacification. Nagai *et.al.*[22, 23] used lanosterol nanoparticles with a particle size distribution from 50 to 400 nm to repair the space and structural collapse in the early stages in the lenses.

They found that it delays the onset of opacification of the lenses with a remarkable lens structure collapse and opacification, but it does not repair them. They speculated that the repeated injection of lanosterol nanoparticles attenuated the manifestation of cataract-related factors and perhaps protects the lenses from oxidative stresses. Shanmugam *et al.*[24] found that 25mM lanosterol in 20% ethanol fails to reverse nuclear opacity of human cataractous nuclei after 6 days of incubation. Daszynski *et al.*[25] failed to repeat the experiments by Zhao *et al.*[12] and Markley *et al*[13] using their same approach. Also Daszynkski *et al.*'s docking simulations shows those two sterols cannot bind to the groove which is formed by the α-crystallin dimer using two wild types (PDB 2WJ7 and 2KLR) and a R120G mutant (PDB 2Y1Z) αB-crystallin.

All efforts were focused on the restoration of cataractous crystallin aggregates, but the effects of lanosterol and 25-hydroxycholesterol on the α-crystallin aggregation process were not investigated. If the compounds could not avoid protein aggregation, they could not solubilize aggregates. Furthermore, the influence of the two sterols on the α-crystallin critical biological function and the chaperone activity were not studied. In this paper, a series of studies were pursued to evaluate whether or not lanosterol and 25-hydroxycholesterol prevented the α-crystallin aggregation and their effects on a native α-crystallin chaperon activity using γ-crystallin as substrate. α and γ-crystallins are highly stable at physiological conditions. The melting temperature of bovine α-crystallin and γ-crystallin is approximately 61°C[26, 27] and 80°C[28, 29], but the turbidity of α-crystallin and γ-crystallin start to increase at around 65°C and 52°C[29]. α-crystallin starts to lose secondary structure at approximately 60 °C[30]. For this study, α-crystallin was incubated at 55 °C. The selection of incubation temperatures was based on the following

considerations: 1) Temperatures above 60°C unfold α-crystallin secondary structures, which changes the potential sterols interaction sites. Therefore, they should be avoided. 2) α-crystallin need long periods of time (up to months) to aggregate at temperatures below 50°C, making the experiments not practical. Thus, 55°C was selected to be the incubation temperature for α-crystallin. Unfortunately, the experiments to evaluate chaperon activity happen too fast at 55 °C (within a week). Then, 50 °C was chosen to evaluate the α-crystallin chaperone activity because at 50 °C, partially unfolded γ-crystallin begins to bind to α-crystallin while the secondary structure of α-crystallin is maintained.[31, 32]

To better understand the interactions, if any, between α-crystallin and sterols, we studied the α-crystallin binding to Cu^{2+} in the presence of lanosterol and 25-hydroxycholesterol. In addition of 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid, dipotassium salt (bis-ANS) fluorescent assay, we explored the interaction between those two sterols and the hydrophobic patches of the α-crystallin. The binding of Cu^{2+} to α-crystallin is critical for α-crystallin. The chaperone activity of α-crystallin has been enhanced upon Cu^{2+} binding[33-35]. The α-crystallin served as a collector of Cu^{2+} and alleviate cytotoxic Cu^{2+} mediated oxidation[31, 34]. Cu^{2+} has been reported to bind to α-crystallins through the mini-αA 70–88 peptide[36], the α-crystallin chaperone active site[10]. The 4-(2-Pyridylazo)resorcinol (PAR) assay has been successfully used to measure Cu^{2+} binding to α-synuclein[37] human αA-crystallin[31]. Therefore, any effect of the sterols on the binding of Cu^{2+} is indirect evidence of them interacting with the α-crystallin chaperone activity site. The bis-ANS is a fluoresce probe which binds to proteins through interaction with the aromatic rings.[38] Ghahramani *et al.*[33], Ghosh

et al.[31] and Raju *et al.*[36] have shown binding of Cu^{2+} to α-crystallin has lower the bis-ANS fluorescence intensity. Raju *et al.* pointed out the bis-ANS binds to the mini αA-crystallin 70-88. And the histamine (H) at position 79 is the binding site for Cu^{2+}.

2. MATERIALS

Fresh bovine lenses were purchased form Animal Technologies, Inc. (Tyler, TX, US) and stored at -20°C in a storage buffer: 1% 2-mercapthanol, 0.245 M acetate buffer at pH 5.0. Bulk lanosterol (55% purity), 25-hydroxycholesterol (>98% purity), 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid, dipotassium salt (bis-ANS) fluorescence probe, 4-(2-Pyridylazo)resorcinol (PAR), GC grade dimethyl sulfoxide (DMSO), guanidine hydrochloride (GdnHCl) and $CuCl_2$ were purchased from Sigma-Aldrich (St. Louis, MO, US). Cu^{2+} standard solution was purchased from Acros (Fisher Scientific, Hampton, NH, US). Ultrafiltered type 1 water (≥18.0 MOhm/cm) was used in all experiments. The other chemicals were analytical grade.

3. METHODS

3.1. BOVINE α-AND γ-CRYSTALLIN ISOLATION

One bovine lens was thawed in water at room temperature and then homogenized in 0.05M Tris pH 7.4 buffer. The homogenate was centrifuged at 15,000 *xg* at 4°C. The γ-crystallin was separated from α- and β-crystallin by running the supernatant through a Sephadex G-75 gel permeate column (GPC) (2.5 x 60 cm) with 0.05M pH 7.4 Tris buffer

as mobile phase. The flow rate was 0.5 ml/min, and the Uv detector was set as 280 nm. The pooled γ-crystallin and the mixture of α- and β-crystallins were dialyzed against reverse osmosis (RO) water for 24 hours with water changes every 8 hours at 4°C. After dialysis, the samples were lyophilized and stored at -20°C. The α-crystallin fraction was separated from β-crystallin with a Sephadex G-200 GPC (2.5 x 60 cm) column using the same buffer as mobile phase, but the flow rate was changed to 0.1 ml/min. The α-crystallin fraction was dialyzed, lyophilized, and stored at -20°C.

3.2. LANOSTEROL ISOLATION AND 25-HYDROXYCHOLESTEROL

The bulk lanosterol was purified by a method from previous work.[39] Briefly, one milliliter of saturated crud lanosterol solution in methanol was injected to an AKTA purifier FPLC (Marlborough, MA, US) equipped with C18 reverse-phase column (4.6 mm x 150 mm). The mobile phase was methanol at a flow rate of 2.0 ml/min. Pure lanosterol was collected and then vacuum-oven dried at 60°C. The 25-hydroxycholesterol was used as 'received' without further purification.

3.3. α-CRYSTALLIN AGGREGATION WITH AND WITHOUT LANOSTEROL AND 25-HYDROXYCHOLESTEROL

The following samples were prepared: α-crystallin only, α-crystallin with 5% DMSO, α-crystallin with 5% DMSO and lanosterol, α-crystallin with 5% DMSO and 25-hydroxycholesterol. The α-crystallin powder was dissolved in 0.1M pH 7.4 phosphate buffer at a concentration of 5.0 mg/ml. Lanosterol and 25-hydroxycholesterol were dissolved in DMSO at 1.0 mg/ml. Fifty microliters of each sterol stock solutions were then mixed with the α-crystallin solution to achieve sterols' final concentrations of 125

μM in the high concentration sterols experiments. In the low sterol concentration experiments, the final sterols concentrations were 0.5 μM and 45 μM for lanosterol and 25-hydroxycholesterol, respectively. For the control samples, 50 μL of DMSO or phosphate buffer were added, instead of the sterol solutions. The samples were transferred to the 2 mL polypropylene test tubes with screw caps, and sealed with Parafilm M (American National Can, Chicago, IL), and then they were incubated on a dry bath at 55°C. The experiments were done in triplicate.

3.4. DYNAMIC LIGHT SCATTERING (DLS)

The particle size distribution of α-crystallin (5.0 mg/ml), lanosterol (0.5μM and 125μM in 5% DMSO), 25-hydroxycholesterol (45μM or 125μM in 5% DMSO), or a mixture of α-crystallin and the two sterols in 0.1M pH 7.4 phosphate buffer were measured using fiber optic quasi elastic light scattering (*FoQels*) (Brookhaven Instruments. Holtsville, NY, US). The instrument operated at 25°C. The laser wavelength was 830 nm, and the detector had a back angle of 135.9°. The samples were put in a glass cylindrical cuvette with a diameter of 1 cm. The correlation function was obtained after scanning for 2 minutes, and it fitted with the CONTIN algorithm. The particle size distribution were calculated by intensity and by number.

3.5. EVALUATION OF α-CRYSTALLIN CHAPERON ACTIVITY WITH γ-CRYSTALLIN WITH LANOSTEROL AND 25-HYDROXYCHOLESTEROL

The following samples were prepared: γ-crystallin, γ-crystallin+ α-crystallin (1:2), γ-crystallin+ α-crystallin (1:2) with 5% DMSO, γ-crystallin+ α-crystallin (1:2) with 5% DMSO and lanosterol, γ-crystallin+ α-crystallin (1:2) with 5% DMSO and 25-

hydroxycholesterol. The α-crystallin and γ-crystallin were dissolved in 0.1M pH 7.4 phosphate buffer at final concentrations of 2.0 mg/ml and 1.0 mg/mL for α-crystallin γ-crystallin, respectively (α : γ 2:1). Fifty microliters of each sterol stock solution was then mixed with α-crystallin- γ-crystallin solutions to obtain sterols at a final concentration of 125 μM in the high concentration sterols experiments. In low sterols concentration experiments, the final sterol concentrations were 0.5 μM and 45 μM for lanosterol and 25-hydroxycholesterol, respectively. For control samples, 50 μL of DMSO or phosphate buffer were added. The samples were transferred to the 2 mL polypropylene test tubes with screw caps, then sealed with Parafilm M and they were incubated on a dry bath at 50°C. The experiments were done in triplicate.

3.6. TURBIDITY

The turbidity of the samples was monitored by a Genesys 5 spectrophotometer (Thermo Fisher Scientific Inc. Waltham, MA, US) at 600 nm. The temperature was controlled by a thermostatic bath at 55°C or 50°C. Samples of 0.8 mL were quickly transferred to the polystyrene disposable semi-micro cuvettes with open tops for turbidity measurements. Blanks were subtracted and averages calculated. The lag-time, t_{lag}, and growth-rate, g_r, were extracted by fitting the equation[40] below to turbidity data:

$$F = \frac{F_{\max}}{(1 + ve^{-k(t-t_m)})^{1/v}} \quad , \qquad (1)$$

where F is the turbidity at 600nm, F_{max} is the maximum turbidity at the steady-state, t_m is the point of maximum growth rate, and v describes the asymmetry of the sigmoid curve.

The growth rate (g_r) is given by $k/(1+v)$ and the lag-time (t_{lag}) is calculated from: $t_m - (1+v)/k$, which is the time where the tangent at t_m crosses the lag-phase baseline.[41]

3.7. MONITORING THE SOLUBLE FRACTION BY HPLC

A Shimadzu HPLC system (Nakagyo-ku, Kyoto, Japan) equipped with a Biobasic 300 size exclusion column was used to monitor the concentration of the soluble fraction. The Uv-vis detector was set at 280 nm. The mobile phase was 0.1 M pH 7.4 phosphate buffer running at 1.0 ml/min. Samples were first centrifuged at 15,000 *xg* at 4°C for 5 minutes, then the supernatant was sampled and diluted. The α-crystallin samples were diluted at 12.5x, and α-crystallin + γ-crystallin samples were diluted at 5x. Ten μL of the samples were injected.

3.8. FOURIER-TRANSFORM INFRARED SPECTROSCOPY (FTIR)

The secondary structure of crystallins was monitored using a Nicolet 6700 FTIR (Thermo Fisher Scientific Inc. Waltham, MA, US) spectrophotometer. Before measurement, the spectrophotometer was purged with dry air, and liquid nitrogen was added to the detector. That was followed by 40 minutes of equilibration. Twenty μL of each sample was dried on a CaF_2 window using a fan for 5~10 minutes. The CaF_2 window was then put in the measuring chamber. A five minute re-equilibration time was given for each time the chamber was opened. The range of 4000 cm^{-1} to 1000 cm^{-1} was recorded at 64 scan with 8 cm^{-1} resolution. Amid I peak, 1700 cm^{-1} to 1600 cm^{-1} of each spectrum was deconvoluted using Origin 2016. The peaks from the deconvolved spectra were assigned to specific secondary structures according to previous studies.[42-44]

3.9. CIRCULAR DICHROISM SPECTROSCOPY (CD)

The 5 mg/ml α-crystallin samples were incubated with 125 μM of lanosterol or 25-hydroxycholesterol at 37 °C for 2 hours. Then, they were loaded on a G-75 GPC column to remove unbound sterols, using 0.1 M pH 7.4 phosphate buffer as the mobile phase and a flow rate of 1 ml/min. The α-crystallin fractions were collected and the concentration was measured at 280 nm with an extinction coefficient of $A^{1\%}_{280}$= 8.42. The concentration of α-crystallin was then adjust to 0.3mg/ml for far UV-CD and 2.0 mg/ml for near UV-CD. The far UV-CD spectra were recorded using a JASCO J-815 spectropolarimeter (Easton,MD, USA) at 25°C in a 0.2 cm path length quartz cell from 200 to 250 nm. The near UV-CD spectra were recorded from 250nm to 340nm with same instrument, but the protein concentration was changed to 2.0 mg/ml and a 1.0 cm path length quartz cell was used. The spectra of proper blanks were subtracted from each protein spectrum. The spectra were analyzed using the BeStSel online sever.[45, 46]

3.10. COPPER ION BINDING OF A-CRYSTALLIN IN THE PRESENCE OF LANOSTEROL AND 25-HYDROXYCHOLESTEROL, 4-(2-PYRIDYLAZO) RESORCINOL (PAR) ASSAY

The PAR assay was adapted from Ghosh *et al.*[31] The excess amount of $CuCl_2$ was added to a 0.1 M pH 7.4 phosphate buffer with 5 mg/ml α-crystallin. Then, 50 μL of DMSO or sterols DMSO solution was added to achieve 125 μM for high concentration of those two sterols, 0.5 μM for low concentration of lanosterol, and 45 μM for low concentration of 25-hydroxycholesterol. After 2 hours of stirring and incubation at 37°C, the samples were centrifuged at 1000 *xg* to sediment the excess amount of $CuCl_2$. The supernatant was loaded on a Sephadex G-75 GPC (2.5 x 30 cm) column to remove

unbound Cu^{2+}, using 0.1 M pH 7.4 phosphate buffer as the mobile phase and a flow rate of 1 ml/min. The α-crystallin fractions were collected and the concentration was measured at 280 nm with an extinction coefficient of $A^{1\%}_{280}$= 8.42. Then, the supernatant was treated for five minutes with 4M GdnHCl to denature the α-crystallin and release the copper ion. After which, 10 μL of a freshly prepared PAR dye solution was added to each sample with a final concentration of the dye at 100 μM. The absorbance at 514nm was recorded and subtracted from the control sample, which had 4M GdnHCl and PAR. Cu^{2+} content was calculated from:

$$Cu^{2+}(\mu M) = 24.33\Delta A_{514nm} + 1.049 \quad (2)$$

Equation 2 was obtained by measuring a series of Cu^{2+} standard solution (Acros Organics) using PAR.

3.11. BIS-ANS FLUORESCENT ASSAY

The samples were prepared by the same procedures described in the PAR assay. The samples after GPC were diluted to 15μM of α-crystallin. That was followed by addition of 10 μL of bis-ANS stock solution (14.8 μM in 95% ethanol) to reach a final bis-ANS concentration of 0.148 μM. The samples were then incubated at 37°C for 20 minutes. The bis-ANS fluorescence was monitored using a Nanodrop 3000 fluorospectrometer (Thermo Fisher Scientific Inc. Waltham, MA, US) with an excitation wavelength of 390 nm and the emission spectrum were recorded from 395 nm to 751 nm. Each sample was measured 9 times and the fluorescence intensity at 490 nm was averaged.

4. RESULTS

4.1. DYNAMIC LIGHT SCATTERING

The solubilities of lanosterol and 25-hydroxycholesterol in 5% DMSO are not available. First we measured the solubility of the two sterols according to Li and Forciniti.[39] They are 0.5 μM for lanosterol and 45 μM for 25-hydroxycholesterol, respectively. Therefore, at a concentration of 125 μM (the high sterols samples used in this study), the high sterols samples used in this study consist of undissolved sterol at equilibrium with a saturated solution. Figures 1 A- D show the particle hydrodynamic diameters distribution by number and by intensity of the two sterols at 125 μM in 5% DMSO phosphate buffer without α-crystallin. The hydrodynamic diameters calculated by number and by intensity were similar. Both sterol "solutions" have large undissolved particles. The hydrodynamic diameter of lanosterol and 25-hydroxycholesterol were around 1500 nm and 3500 nm. Figures 1 I-L showed the low concentration sterols samples with α-crystallin. The hydrodynamic diameters calculated by number and by intensity were similar. The only detectable particle is α-crystallin, which has a hydrodynamic diameter around 10nm, which was the same of α-crystallin (data not shown). Figures 1E - H showed that by adding α-crystallin to high concentration sterols solutions, the particle size distributions calculated by number and by intensity were different. The particle size distribution calculated by intensity showed that particle diameter of undissolved 25-hydroxycholesterol (Figure 1H) decreased to approximately 600 nm, however, the undissolved lanosterol particle diameter (Figure 1F) barely changed. The particle diameter distribution of lanosterol with α-crystallin calculated by

number was approximately 120 nm and for 25-hydroxycholesterol, which was approximately 17 nm (Figure 1 E and G). Those results suggested that there were two sets of particles in the samples: 1) one set of small particles with a large population that correspond to α-crystallin and 2) a set of large particles with a small population that corresponds to the undissolved sterols. The hydrodynamic diameter of α-crystallin increased upon addition of high concentration of the sterols, which suggests that α-crystallin forms large aggregates in the presence of undissolved sterols particles.

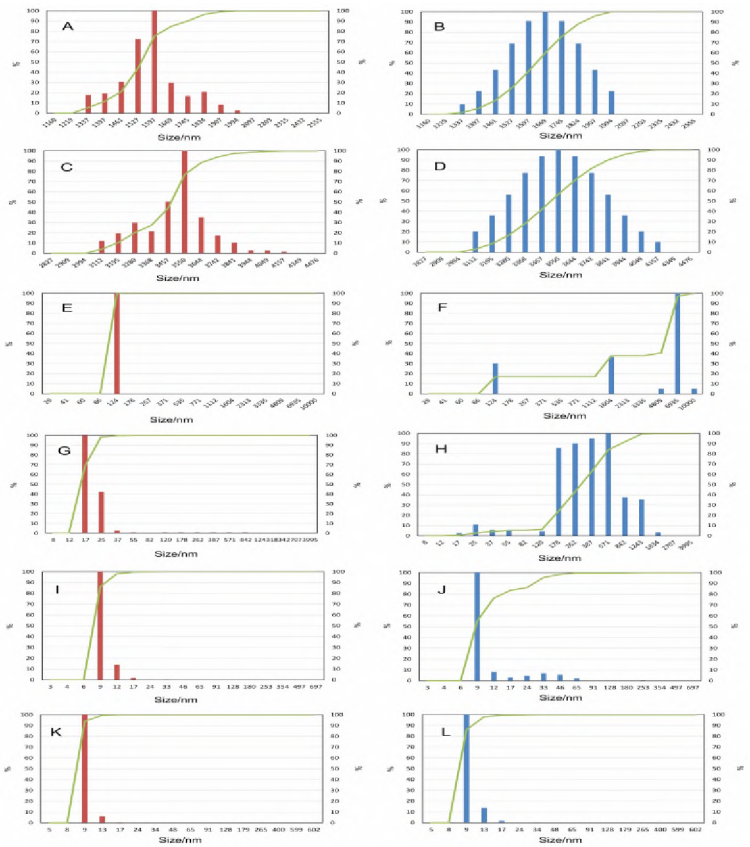

Figure 1. Dynamic light scattering (DLS) of lanosterol and 25-hydroxycholesterol in 5% DMSO phosphate buffer 0.1M pH 7.4. Left panel: by number; Right panel: by intensity. A and B, 125uM lanosterol. C and D, 125μM 25-hydroxycholesterol. E and F, 5mg/mL α-crystallin+ 125μM lanosterol. G and H, 5mg/mL α-crystallin+ 125μM 25-hydroxycholesterol. I and J, 5mg/mL α-crystallin+ 0.5μM lanosterol. K and L, 5mg/mL α-crystallin+ 45μM 25-hydroxycholesterol

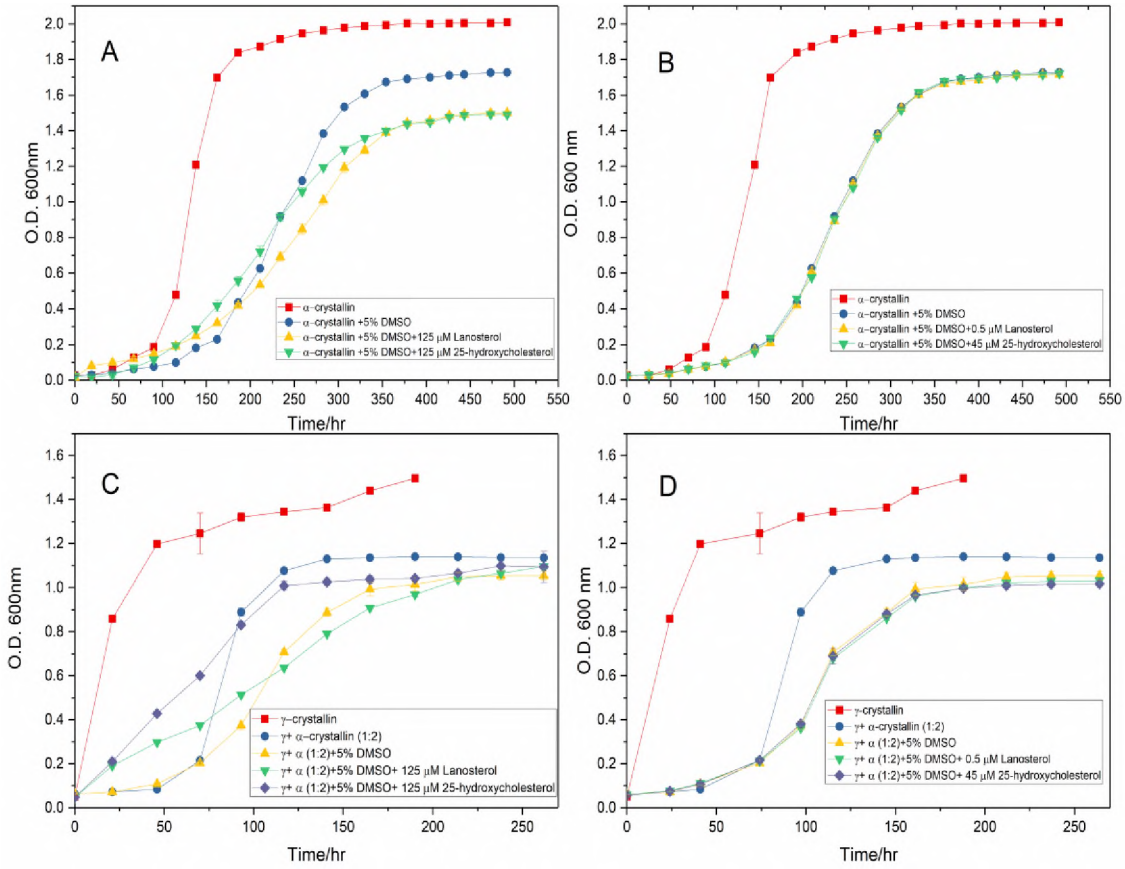

Figure 2. Turbidity of α-crystallin or α-crystallin and γ-crystallin mixture in 0.1M pH 7.2 phosphate buffer. A, turbidity of α-crystallin incubated with 125μM sterols at 55°C. B, turbidity of α-crystallin incubated with 0.5μM lanosterol or 45μM 25-hydroxycholesterol at 55 °C. C, turbidity of α-crystallin and γ-crystallin incubated with 125μM sterols at 50 °C. D, turbidity of α-crystallin and γ-crystallin incubated with 0.5μM lanosterol or 45μM 25-hydroxycholesterol at 50 °C

Figures 2 C and D show the aggregation kinetics of α-crystallin and γ-crystallin mixtures. The lag-times and growth rates are summarized in Figure 3 B and D. The samples that only had γ-crystallin aggregated within half an hour at 50 oC. The samples that had α-crystallin yielded longer lag-times and slower growth rates, which demonstrated the α-crystallin chaperone activity. Compared to the samples of α-crystallin and γ-crystallin without DMSO, the samples with 5% DMSO had longer lag-times and

slower growth rates, which again may be due to the preferential hydration of the protein caused by DMSO. The lag-times for the samples with high concentration of both sterols decreased drastically, but the growth rates were slower than in the DMSO control. The lag-times and growth rates for low sterol concentration samples were not significantly different from the DMSO controls.

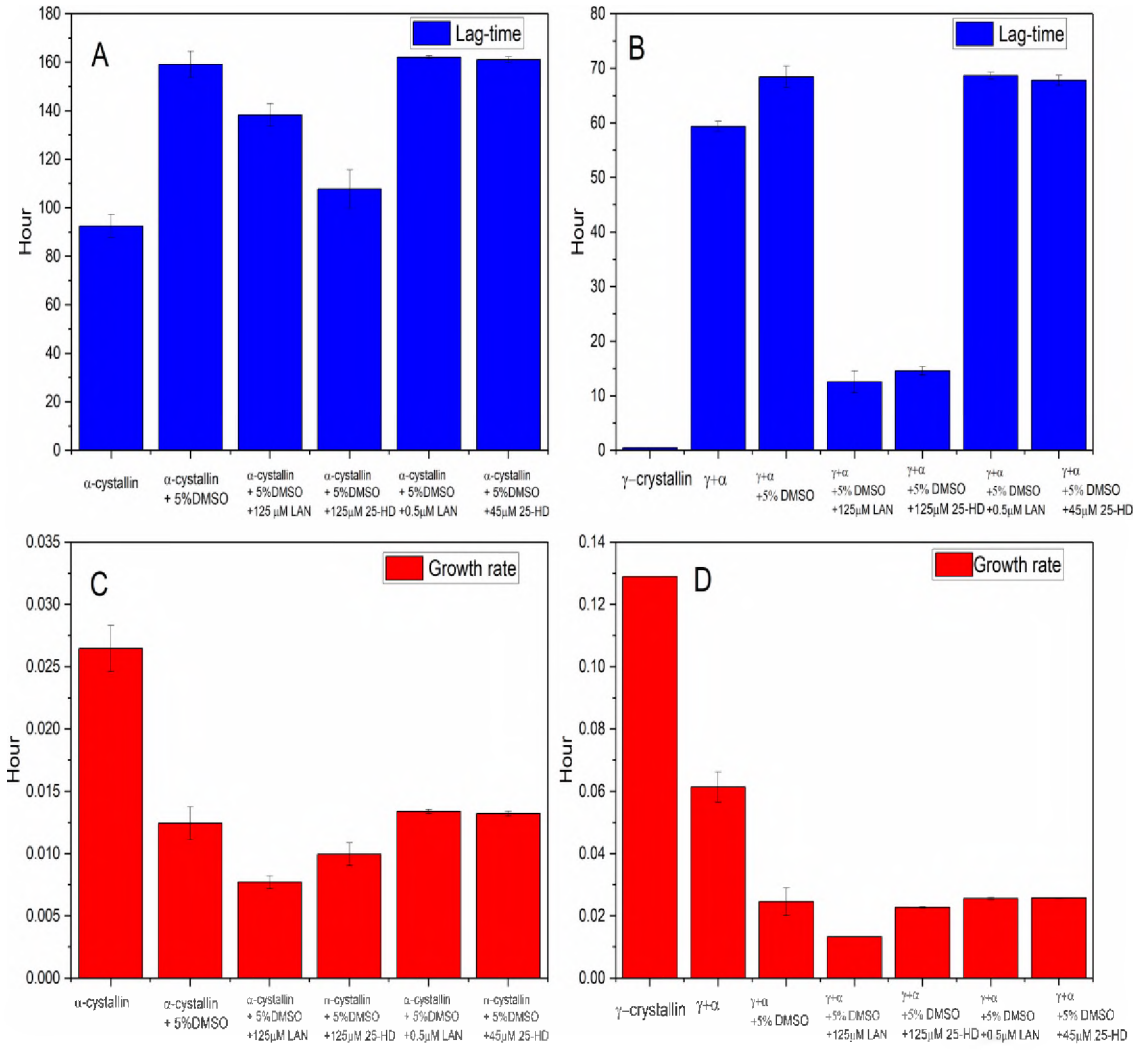

Figure 3.Lag-time and growth rate of α-crystallin or α-crystallin and γ-crystallin mixture aggregation kinetics. A and C, lag-time and growth rate of α-crystallin. B and D, lag-time and growth rate of α-crystallin and γ-crystallin mixture

4.2. SOLUBLE FRACTION

The soluble fraction of α-crystallin incubated with high concentrations of the two sterols was plotted together with turbidity data (Figure 4A). The soluble fraction of α-crystallin decreases as the turbidity increases. The soluble fraction of samples without DMSO decreases faster than the samples with DMSO as observed in the turbidity measurements. The retention times of α-crystallin (Figure 4B) were shorter upon heating, which implies that the apparent molecular weight of α-crystallin has increased. The increase of α-crystallin apparent molecular weight upon heating has been reported by Putilina *et.al.*[50] who used gel filtration and small angle x-ray scattering (SAXS). That increase in the apparent molecular weight α-crystallin may also be the reason for the decrease in the soluble fraction at the beginning of heating. The decrease in the soluble fraction began at around 24 hrs when the turbidity was still in the lag-phase may reflect the lost of soluble aggregates during centrifugation prior to HPLC analysis. At around ~250 hrs, there is an inverse correlation between the amount of the soluble fraction and turbidity values, i.e., the soluble fractions of the sample with DMSO are in the order of α-crystallin+ 125μM 25-hydroxycholesterol > α-crystallin > α-crystallin + 125μM lanosterol, but the order of for the turbidity values is inversed.

Figure 5A shows that the soluble fraction of α-crystallin incubated with low concentration of the two sterols. In this case, the soluble fraction of α-crystallin was not significantly different from the control sample (α-crystallin in 5% DMSO). Figure 5B shows that the retention times of the α-crystallin were similar to the 5% DMSO control.

Figure 6A-C shows the soluble fraction of the α-crystallin and γ-crystallin mixtures incubated with high concentration of the sterols. The total soluble fraction

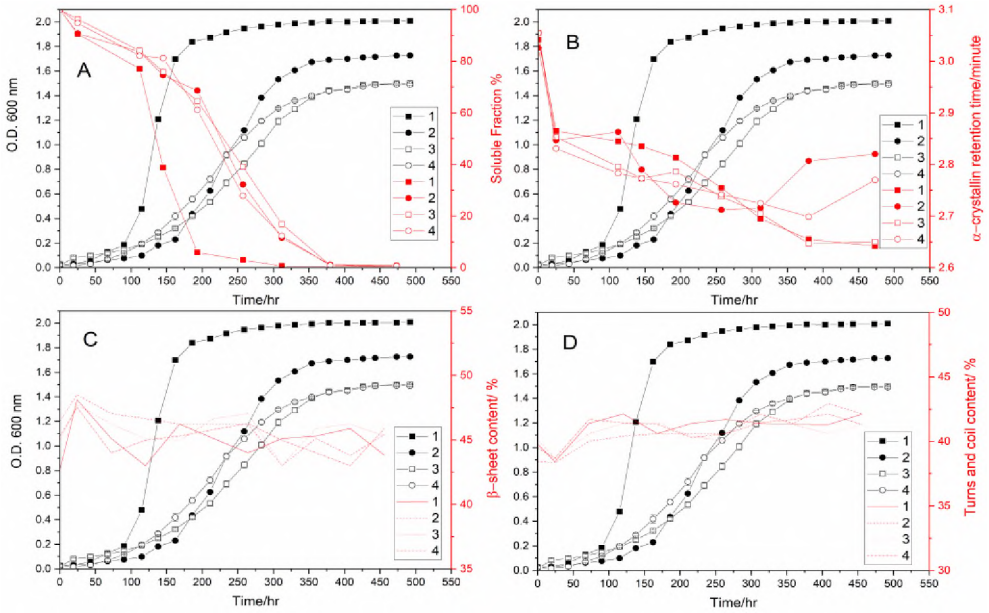

Figure 4. The α-crystallin incubated with 125μM lanosterol or 25-hydroxycholesterol. 1, α-crystallin control. 2, α-crystallin 5% DMSO control. 3, α-crystallin+ 125μM lanosterol in 5% DMSO. 4, α-crystallin+ 125μM 25-hydroxycholesterol in 5% DMSO. A, soluble fraction of α-crystallin. B, retention time of α-crystallin. C, β-sheet content. D, turns and coil content

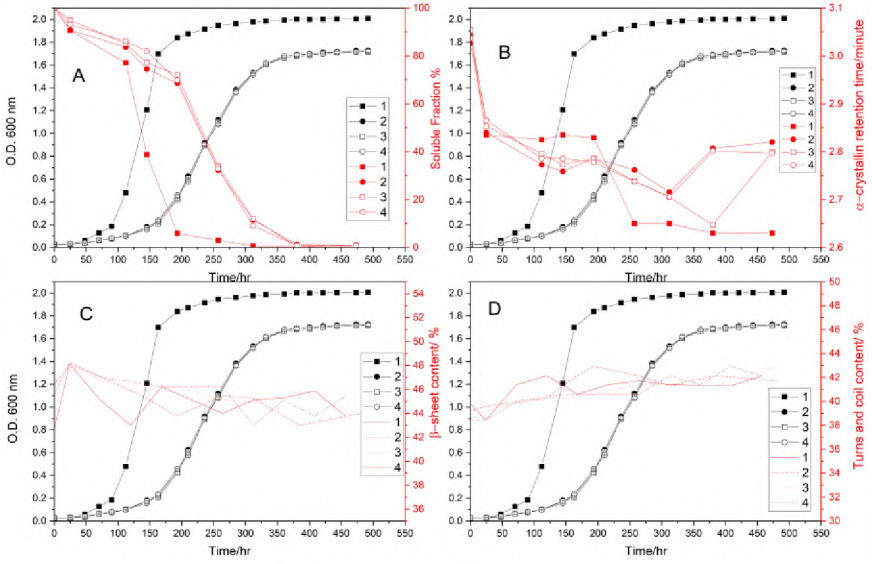

Figure 5. The α-crystallin incubated with 0.5μM lanosterol or 45μM 25-hydroxycholesterol. 1, α-crystallin control. 2, α-crystallin 5% DMSO control. 3, α-crystallin+ 0.5μM lanosterol in 5% DMSO. 4, α-crystallin+ 45μM 25-hydroxycholesterol in 5% DMSO. A, soluble fraction of α-crystallin. B, retention time of α-crystallin. C, β-sheet content. D, turns and coil content

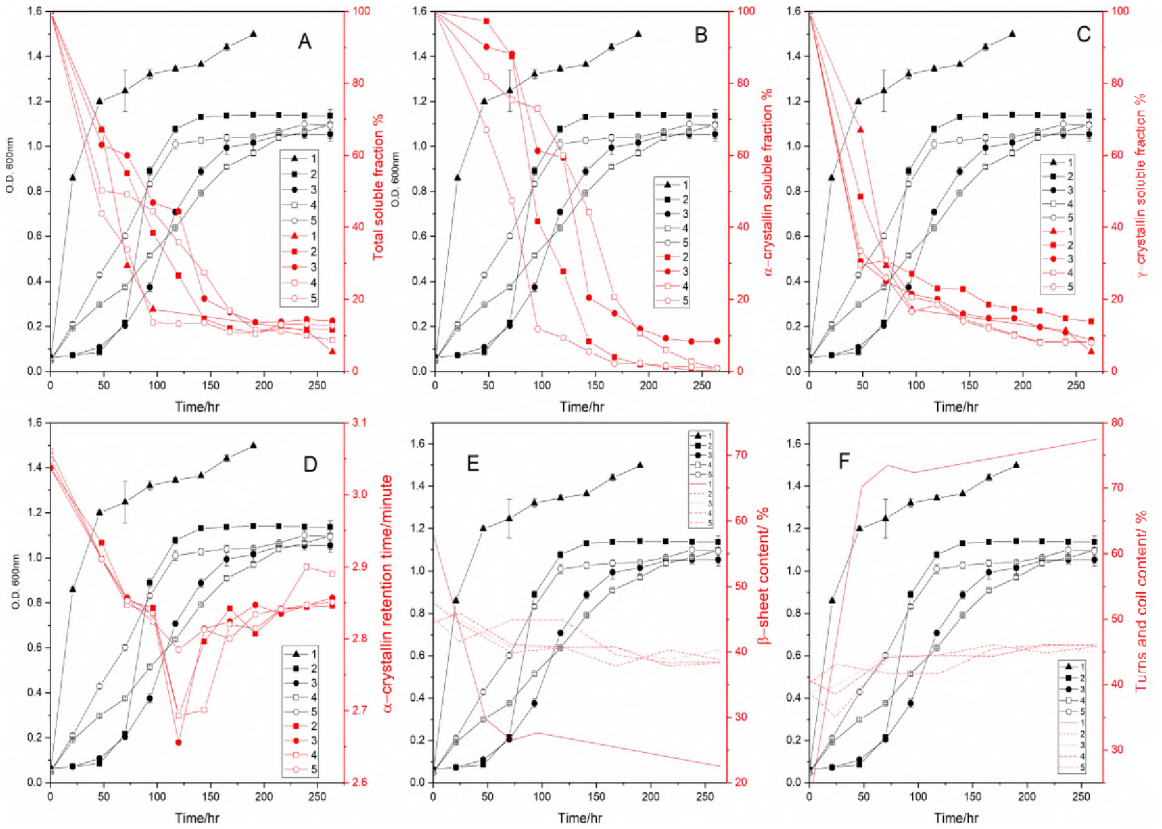

Figure 6. The α-crystallin and γ-crystallin mixture incubated with 125μM lanosterol or 25-hydroxycholesterol. 1, γ-crystallin. 2, α-crystallin and γ-crystallin mixture (2:1) control. 3, α-crystallin and γ-crystallin mixture (2:1) 5% DMSO control. 4, α-crystallin and γ-crystallin mixture (2:1)+ 125μM lanosterol in 5% DMSO. 5, α-crystallin and γ-crystallin mixture (2:1)+ 125μM 25-hydroxycholesterol in 5% DMSO. A, total soluble fraction. B α-crystallin soluble fraction. C, γ-crystallin soluble fraction. D, retention time of α-crystallin. E, β-sheet content. F, turns and coil content

includes both soluble α-crystallin and γ-crystallin, and it decreased as the turbidity increases (Figure 6A). The soluble fraction data are compared against turbidity in Figures 6B and C. The γ-crystallin soluble fraction drastically decreases within 50 hours, but the turbidity of the samples with α-crystallin is still in the lag phase, which demonstrates α-crystallin chaperone activity. Putilina and co-workers[50] also showed that α-crystallin prevent γ-crystallin from aggregation. The γ-crystallin soluble fraction decreased faster in

the samples with 5% DMSO than the samples without DMSO (Figure 6C); i.e., at approximately 50 hours, the α+γ without DMSO had around 50% γ-crystallin soluble fraction and the samples with DMSO had around 30% γ-crystallin soluble fraction. Therefore, DMSO promotes unfolding of the γ-crystallin. The unfold γ-crystallin bound to α-crystallin to form α/γ complex within 9 hours at 55 ^{o}C.31 Therefore, the soluble fraction of α-crystallin included α-crystallin and α/γ complex. In Figure 6C, the soluble fraction of γ-crystallin with high concentration of sterols decreased to 30% within 50 hours. The turbidity is also increased but not to the extent of the samples containing only γ-crystallin. Figure 6B shows that the α-crystallin soluble fraction also decreased in the samples containing high concentrations of sterols. Therefore, the increase in turbidity is mainly caused by the loss of α-crystallin or the α/γ complex. Compared to the α-crystallin turbidity data, the α/γ complex has a higher aggregation propensity. In spite of the fact that α-crystallin was incubated at higher temperatures (55^{o}C vs 50^{o}C) and higher concentrations (5 mg/mL vs 3 mg/mL). Figure 6D shows that the retention time of α-crystallin first decreases to a minimum, and then increased. The decrease in α-crystallin retention time indicates that the apparent molecular weight of α-crystallin increases by forming higher oligomers and by binding to unfolded γ-crystallin.[50] The minimum retention time appeared at the end of the growth phase of each sample. It indicated there was remaining soluble low molecular weight α-crystallin in the solution.

Figure 7A-C show the soluble fraction of α-crystallin and γ-crystallin incubated with low concentration of lanosterol or 25-hydroxycholesterol. Figure 7D shows the retention time of α-crystallin vs incubation time. No major differences were found between samples containing sterols and DMSO controls.

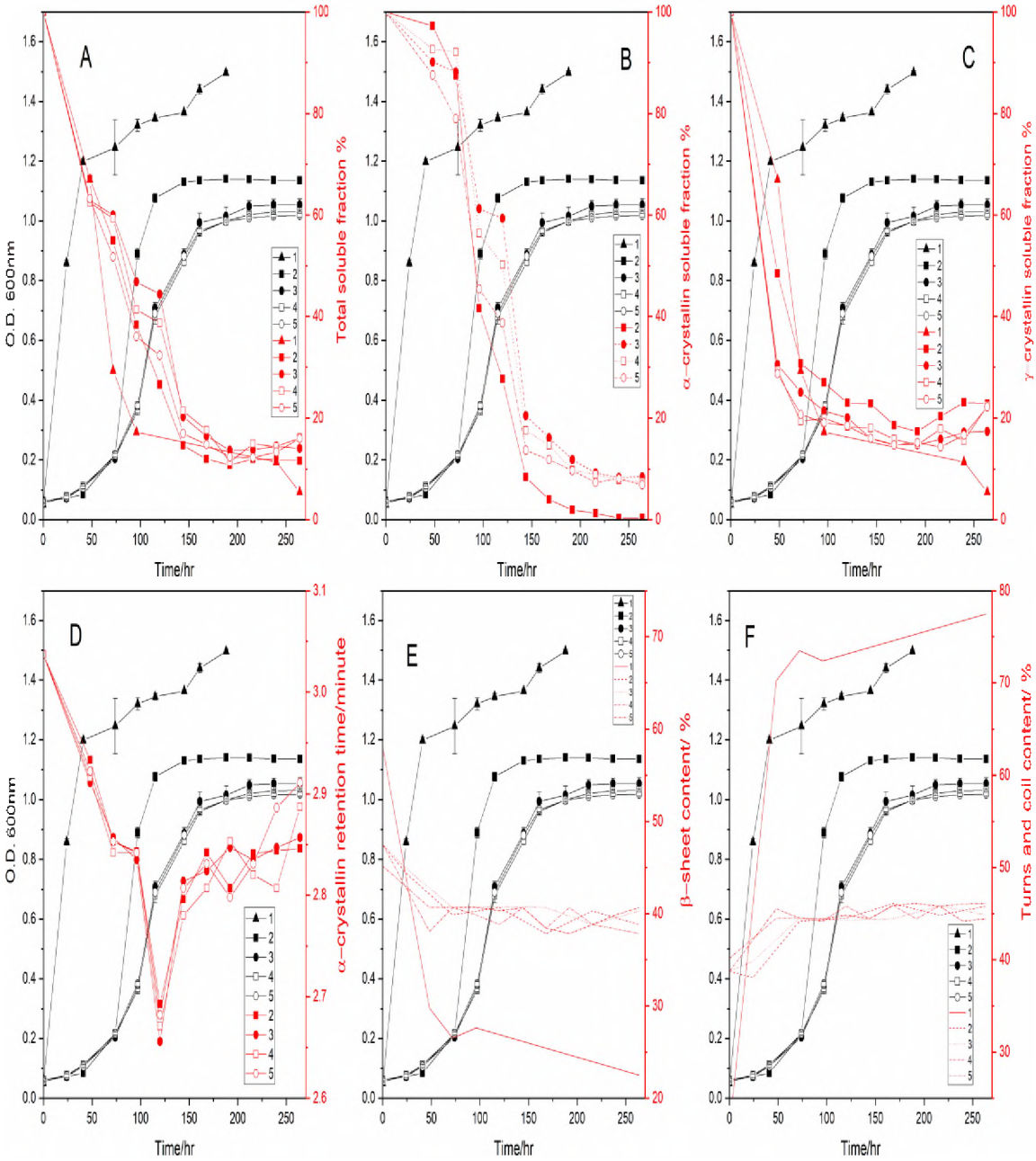

Figure 7. The α-crystallin and γ-crystallin mixture incubated with 0.5μM lanosterol or 45μM 25-hydroxycholesterol. 1, γ-crystallin. 2, α-crystallin and γ-crystallin mixture (2:1) control. 3, α-crystallin and γ-crystallin mixture (2:1) 5% DMSO control. 4, α-crystallin and γ-crystallin mixture (2:1)+ 0.5μM lanosterol in 5% DMSO. 5, α-crystallin and γ-crystallin mixture (2:1)+ 45μM 25-hydroxycholesterol in 5% DMSO. A, total soluble fraction. B α-crystallin soluble fraction. C, γ-crystallin soluble fraction. D, retention time of α-crystallin. E, β-sheet content. F, turns and coil content

4.3. SECONDARY STRUCTURES MONITORED BY FTIR

The protein secondary structures were monitored by FTIR through the aggregation processes. The Amide I peaks were deconvoluted. Figure 8 shows an example of the deconvolution of the Amide I peak, which were at the beginning and at the end of the α-crystallin that was incubated with 125 μM lanosterol at 55°C. The deconvoluted peaks were assigned to the following secondary structures: 1630 and 1691 cm^{-1} are β-sheet, 1641 cm^{-1} is α-helix, 1661 cm^{-1} is turns and 1676 $cm^{-1}$is random coil.[42-44] The area percentages for each deconvoluted peak were used to represent the content of each secondary structure.

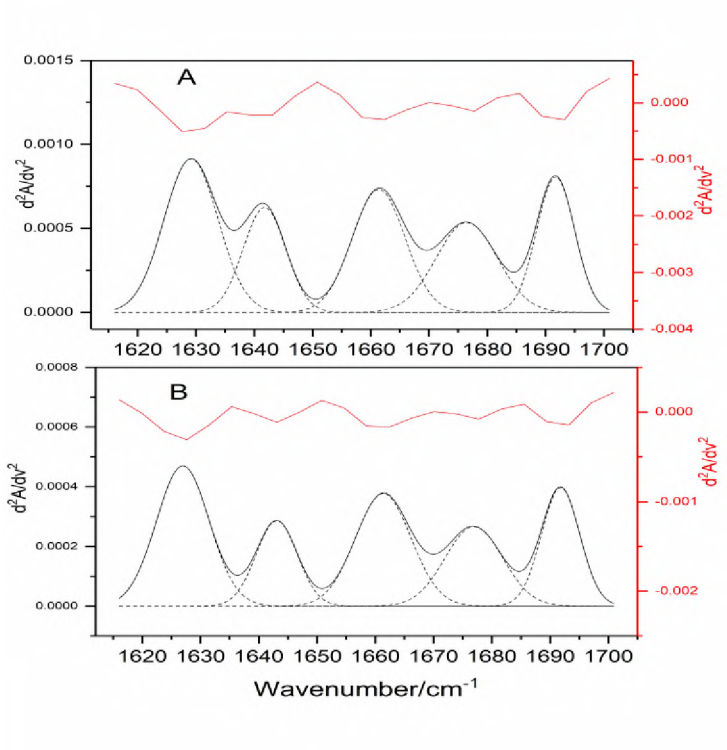

Figure 8. Examples of FTIR Amid I spectra deconvolution. The red lines are second derivative of Amid I peak. The black slid line is the sums of the deconvoluted peaks, and the dash line represent the deconvoluted peaks. The peaks are: 1630 and 1691 cm^{-1} are β-sheet, 1641 cm^{-1} is α-helix, 1661 and 1676 cm^{-1} are turns and random coil. A, α-crystallin + 125μM lanosterol before incubated at 55°C. B, α-crystallin + 125μM lanosterol incubated at 55°C for 460 hours

Figure 4 C and D and Figure 5 C and D show the percentages of β-sheet and turns and coil of α-crystallin *vs* incubation time. The native α-crystallin secondary structure contents were α-helix: 15%, β-sheet: 45% and turns and random coil: 40%, which agree with previous publications.[30, 47] The β-sheet decreased slightly from 45% to 42% during the incubation. The percentages of the turns and random coil increased from 39% to 42%. The α-helix stayed at approximately 15% throughout incubation. No significant changes to the secondary structure of α-crystallin at 55°C were expected because previous tests showed the significantly loss of α-crystallin secondary structures usually been seen above 60°C.[30] The changes to α-crystallin secondary structure of the samples containing both sterols were similar to the DMSO controls, regardless of high or low sterol concentration.

Figure 6 E and F and Figure 7 E and F show β-sheet and turns and random coil percentages verse incubation time of α-crystallin and γ-crystallin samples. The deconvoluted peaks of native γ-crystallin were: 1638 cm^{-1} and 1690 cm^{-1} (β-sheet, 58%), 1661 cm^{-1} (α-helix 20%) and 1675 cm^{-1} (turns 21%). The secondary structure contents of the γ-crystallin measured in this study were similar to previous publication.[48] After incubating for 24 hours, a set of new peaks were found: 1618 and 1634 were assigned to β-sheet (30%), 1651 was assigned to random coil (32%), 1670 and 1682 were assigned to turns (39%). At the end of incubation (260 hours), the α-helix of γ-crystallin was completely lost, the percentage of β-sheet content was reduced to 22%, and the turns and coil contents were raised to 78%. In the samples containing both α-crystallin and γ-crystallin, the infrared signal was a combination of both crystallins. The β-sheet content was reduced from 47% to 38%, and the turns and coil contents were increased from 39%

to 46%; the α-helix still held at approximately 15%. This study argued that these changes were caused by the unfolding of gamma-crystallin. The secondary structure contents of α-crystallin and γ-crystallin had no significant differences between the samples containing lanosterol and 25-hydroxycholesterol and the DMSO controls.

4.4. Cu^{2+} CONTENT OF α-CRYSTALLIN

Figure 9 shows the Cu^{2+} content of α-crystallin with and without the two sterols measured by PAR assay. The α-crystallin isolated for this study in the lab had no Cu^{2+} content (data not shown). Figure 9 shows that neither high nor low concentrations of the two sterols had effect on the Cu^{2+} binding capacity of α-crystallin. The α-crystallin bound to Cu^{2+} at a ratio of 1:3, which was lower than the mini αA-crystallin 70-88 that bound to Cu^{2+} at a ratio of 1:1[36], and the recombined human αA-crystallin bound at (1:2).[31] The Cu2+ binding capacity of α-crystallin oligomers is lower than the one for the mini αA-crystallin monomer,[36] and the one for recombinant αA-crystallin since the recombinant protein forms oligomers of different size.[49] And the incubation time and temperature may affect the binding. The mini αA-crystallin incubated with Cu^{2+} for 30 minutes at 25 °C and recombinant αA-crystallin incubated with Cu^{2+} for 9 hours at 55 °C.

Figure 10A shows the fluorescence spectra of bis-ANS. The fluorescence intensity of 490nm of each sample is summarized in Figure 10B. The results were sorted into two groups: with Cu^{2+} and without Cu^{2+}. Adding Cu^{2+} lowered the bis-ANS fluorescence intensity.The mini αA-crystallin region 70–88 KFVIFLDVKHFSPEDLTVK bound Cu^{2+} through 79 His[36] that was negatively charged at pH 7.4. The 71 Phe, 74 Phe and 80 Phe may interact with bis-ANS through the

aromatic side chains[38]. The Cu^{2+} and bis-ANS probe competed for the binding sites[36]. The lanosterol and 25-hydroxycholesterol had no significant effect on the bis-ANS fluoresce intensity, regardless of high or low sterol concentration.

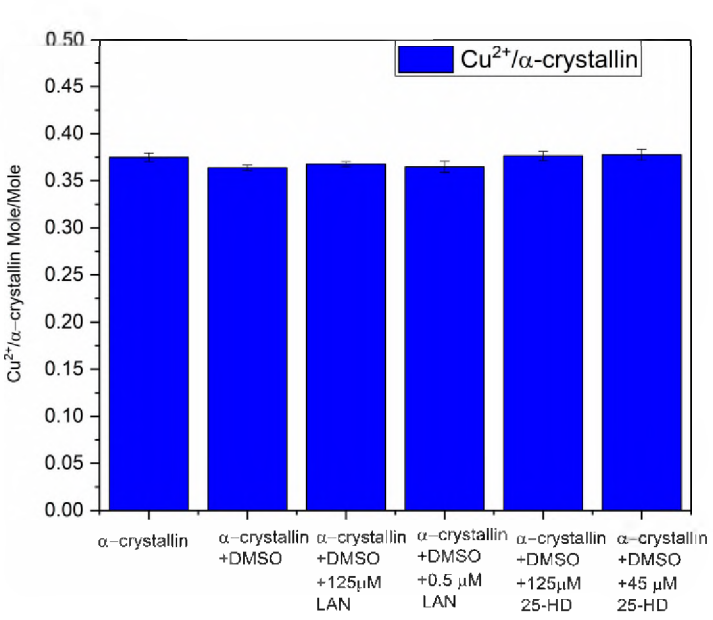

Figure 9. The Cu^{2+} content of α-crystallin. The α-crystallin incubated with access amount of Cu^{2+}, then separated by GPC. The Cu^{2+} content was measured by PAR assay. 25-HD represent the 25-hydroxycholesterol

4.5. CIRCULAR DICHROISM

To corroborate the FTIR-based protein secondary structures, far-Uv CD measurements were used (Figure 11). The α-crystallin incubated with 125μM lanosterol or with 25-hydroxycholesterol was loaded on a G-75 GPC to remove particles that interfered with the CD measurements. As shown in the Figure 11B, the secondary structures of samples incubated with sterols and of DMSO control samples showed no significant differences. Therefore, the secondary structure contents were similar to the results of FTIR. In the near-Uv CD spectra, shown in Figure 12. The tertiary structure of

the samples with the two sterols had no significant difference compared to the DMSO control samples.

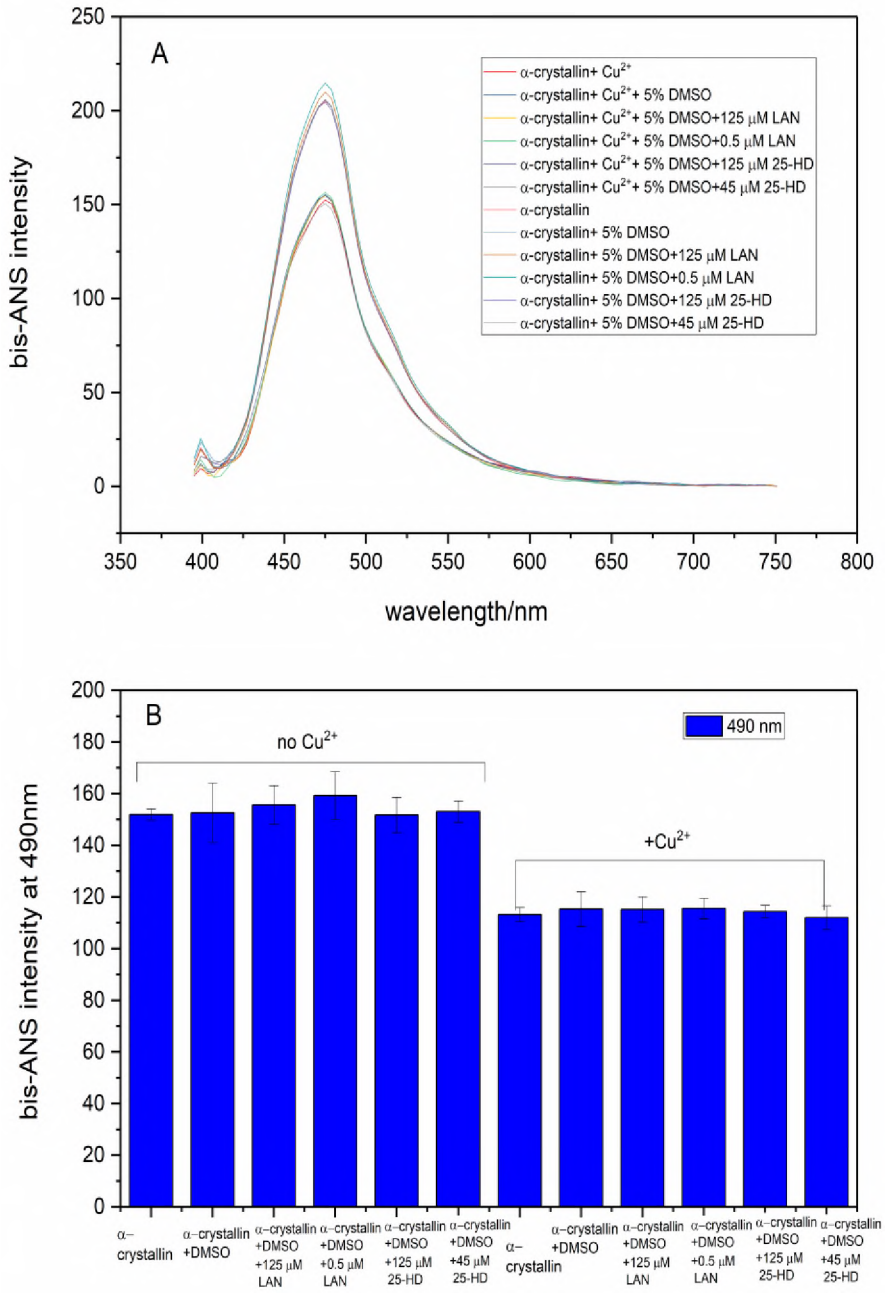

Figure 10. bis-ANS fluorescence spectra of α-crystallin. A, bis-ANS fluorescence spectra of 15μM α-crystallin. The α-crystallin incubated with sterols or controls for 2 hours then separated by GPC. The blank of 30μM bis-ANS in phosphate buffer was subtracted. B, the fluoresce intensity at 490 nm

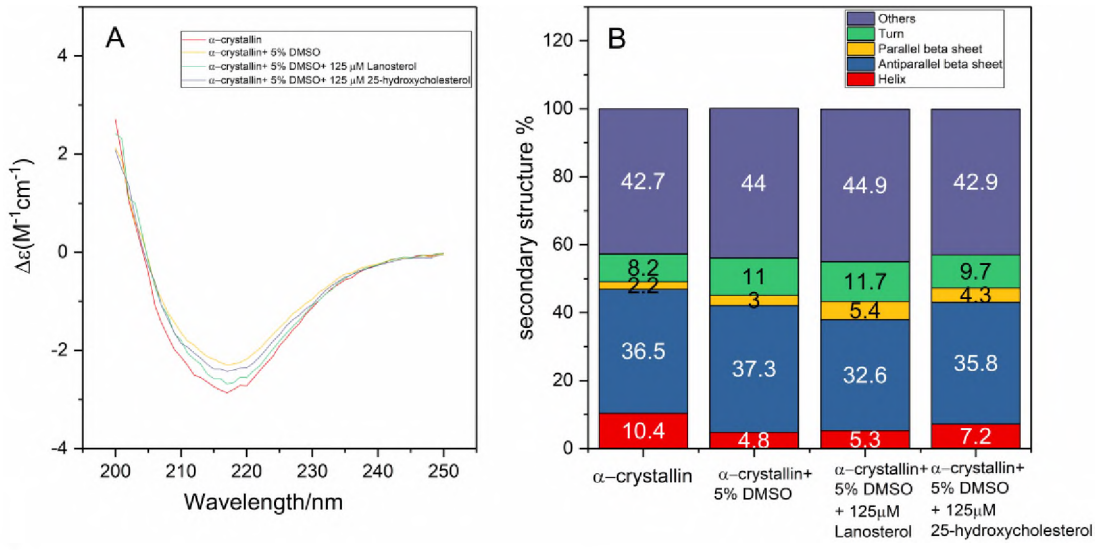

Figure 11. Far-UV CD spectra of the α-crystallin after GPC. A, Far-UV CD spectra. B, secondary structure of α-crystallin which are analyzed by BeStSel online sever.[45]

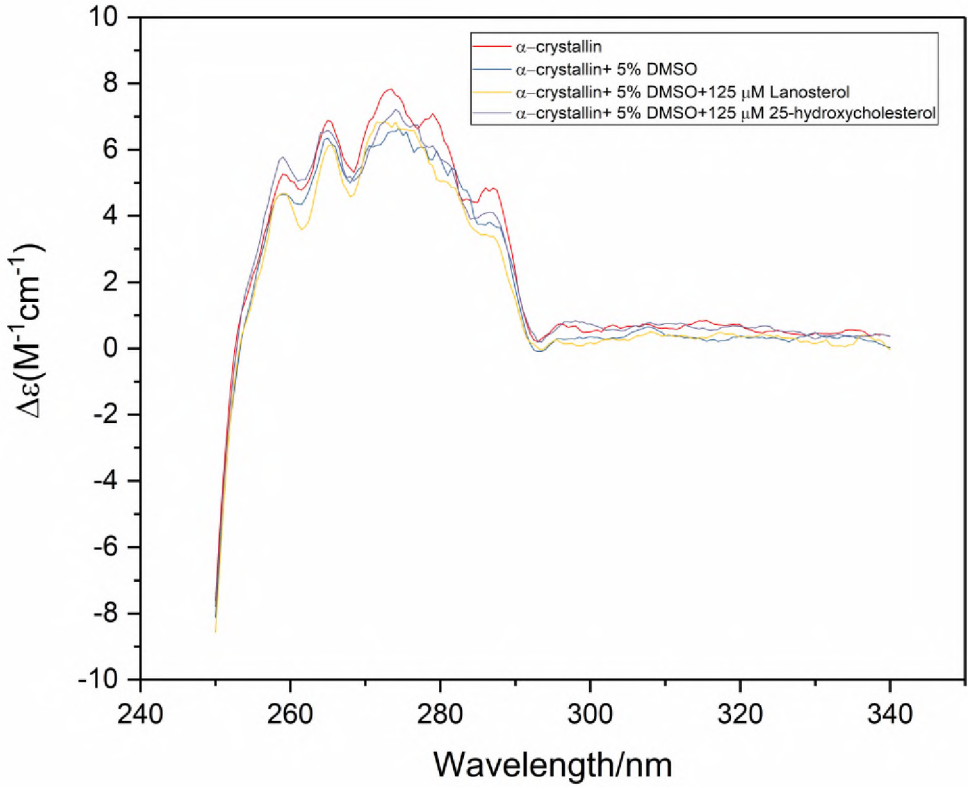

Figure 12. Near-UV CD spectra of α-crystallin with 125 μM lanosterol or 125 μM 25-hydroxycholesterol and controls

5. DISCUSSIONS

α-crystallin has internal cavity structures,[50, 51] which may trap small molecules,[52] such as dexamethasone,[52] 1-anilinonaphthalene-8-sulfonic acid (ANS)[53] and acrylamide.[54] Augusteyn *et al.*[52] concluded that this phenomenon was a nonspecific partitioning. By analogy, the solubilization of 25-hydroxycholesterol by α-crystallin was reported by Puttur *et al.*,[25] who explained the solubilization by speculating that 25-hydroxycholesterol was trapped by α-crystallin oligomers, but not lanosterol. It is not known why lanosterol does not partition into the same cavities. Our DLS measurements show that the undissolved particle size of 25-hydroxycholesterol was reduced from 3500 nm to 600 nm by adding α-crystallin (Figures 1 D and H). On the contrary, the undissolved particle size of lanosterol was not significantly affected by the addition of the α-crystallin (Figures 1 B and F). Therefore, reduction of the undissolved 25-hydroxycholesterol particle size may be due to the nonspecific partitioning whereas lanosterol does not partitioned in agreement with Puttur *et al*'s hypothesis.

Lanosterol and 25-hydroxycholesterol failed to prevent α-crystallin aggregation. Moreover, at sterol concentrations of 125 μM, α-crystallin aggregation was promoted. The turbidity of α-crystallin incubated with low concentration of the two sterols, showing that the sterols do not change the aggregation kinetics. The α-crystallin chaperone activity was not affected by low concentrations of the two sterols. This indicated that completely dissolved lanosterol or 25-hydroxycholesterol did not interact with α-crystallin. In the turbidity data with high concentration of lanosterol or 25-hydroxycholesterol, the lag-times were shorter than the DMSO control samples, which implied that the solid sterol

particles served as nucleation sites. Based on the DLS measurements, the solutions containing high concentration sterols had large undissolved particles that may have affected the aggregation kinetics.

The α-crystallin chaperone activity was evaluated by using γ-crystallin as substrate. The lag-times for α+ γ-crystallin samples with high concentration of the two sterols were shorter than the time yielded by the DMSO control. The undissolved sterol particles altered the aggregation kinetics. Figure 6B and C show that the increase in turbidity was caused by α/γ complex or α-crystallin precipitation. Because the turbidity of the high concentration sterols samples did not increase as fast as the γ-crystallin control samples, so it was concluded that increase in turbidity was due to decreases in the soluble fraction of α-crystallin and α/γ complex. The two sterols promoted the aggregation of α/γ complex and α-crystallin. The chaperone activity of α-crystallin was not affected by the lanosterol or 25-hydroxycholesterol. This conclusion was further supported by the PAR and bis-ANS assay that the chaperone activity sites were not affected by the two sterols, regardless of high or low concentrations of the two sterols.

The α-crystallin samples incubated with high concentration of the two sterols were selected to be assessed by CD measurements because high concentration of sterols changed the aggregation kinetics. The 125 μM of lanosterol or 25-hydroxycholesterol did not change the secondary structures of α-crystallin, which was confirmed by far UV-CD and FTIR. The near UV-CD spectra showed that the tertiary structures of α-crystallin were not affected by high concentration of the two sterols either.

The PAR assay showed the Cu^{2+} content of α-crystallin was not affected regardless of the sterol concentration used. Therefore, the chaperone binding sites of

alpha-crystallin were not occupied by the sterols. However, the binding to other sites cannot be excluded, but that can be addressed by looking at the bis-ANS assays. The bis-ANS bound to hydrophobic sites on the α-crystallin that included the Cu^{2+} binding site. The bis-ANS fluorescence intensity of samples incubated with Cu^{2+} was not affected by the two sterols. That confirmed the result of the PAR assay that stated that the two sterols did not affect the Cu^{2+} binding capacity of α-crystallin. The bis-ANS fluorescence intensity of the samples incubated without Cu^{2+} was not affected by the two sterols in any concentration. It can be stated that the sterols did not bind to the chaperone site. The bis-ANS and the Cu^{2+} competed for the chaperone binding site, which was the mini αA-crystallin region. The affinity of bis-ANS was weaker than that of Cu^{2+}. The possibility of the sterols having weaker interactions than the bis-ANS with the chaperone binding site was not ruled out by this study.

The α-crystallin dimer interface was another potential binding site for sterols beyond the chaperone binding site. Makley *et al.*[13] and Daszynski *et al.*[25] studied the possible interactions between 25-hydroxycholesterol or lanosterol and the α-crystallin dimer interface by docking simulations. Both of them predicted the K_d (dissociation constant that is the ligand concentration at which half the protein molecules will have a ligand bound) of the two sterols were in the high micromolar or even millimolar range; i.e. K_d values for lanosterol-2KLR (wildtype αB-crystallin) and 25-hydroxycholesterol dimer interface were 73.63 mM 25-hydroxycholesterol and 1.22 mM, respectively.[25] Daszynski *et al.*[25] concluded that this high concentration of the two sterols could not be achieved clinically in the lens. Based the experimental results of this study and the conclusion from Daszynski *et al.*, the interactions, if any, between lanosterol or 25-

hydroxycholesterol and α-crystallin are not specific. The presence of undissolved sterols particles that served as aggregation nuclei may be the reason that the undissolved sterols promoted α-crystallin and α/γ complex aggregation.

6. CONCLUSIONS

In this study, it was found that lanosterol and 25-hydroxycholesterol failed to prevent α-crystallin aggregation induced by heating. Sterol concentration of 125μM promoted the aggregation of α-crystallin and of the α/γ complex possibly by serving as nucleation sites. The chaperone activity of α-crystallin was not affected by the two sterols, which suggested that the sterols did not bind to the chaperone binding site. This was confirmed by the bis-ANS results. FTIR and far-UV CD analysis showed that the secondary and tertiary structures of α-crystallin were not affected by the presence of the two sterols, regardless of their concentrations. Furthermore, the two sterols had no significant effect on the Cu^{2+} or on the bis-ANS binding capacity of α-crystallin. Therefore, hydrophobic sites were not interaction sites for the two sterols and α-crystallin. Thus, no evidence was been found that the two sterols interacted specifically with the α-crystallin.

REFERENCES

[1] National Eye Institute. Cataract Data and Statistics, 2019.

[2] Moreau, K. L., King, J. A., Protein Misfolding and Aggregation in Cataract Disease and Prospects for Prevention. *Trends in molecular medicine* 2012, *18*, 273-282.

[3] Michael, R., van Marle, J., Vrensen, G. F. J. M., van den Berg, T. J. T. P., Changes in the refractive index of lens fibre membranes during maturation – impact on lens transparency. *Experimental Eye Research* 2003, *77*, 93-99.

[4] Ray, N. J., Biophysical chemistry of the ageing eye lens. *Biophys Rev* 2015, *7*, 353-368.

[5] Truscott, R. J., Friedrich, M. G., The etiology of human age-related cataract. Proteins don't last forever. *Biochimica et biophysica acta* 2016, *1860*, 192-198.

[6] Xiangjia Zhu, A. K., Roger J.W. Truscott, Age-Dependent Denaturation of Enzymes in the Human Lens: A Paradigm for Organismic Aging? *Rejuvenation Research* 2010, *13*, 553-560.

[7] Laganowsky, A., Benesch, J. L., Landau, M., Ding, L., *et al.*, Crystal structures of truncated alphaA and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. *Protein science : a publication of the Protein Society* 2010, *19*, 1031-1043.

[8] Bagnéris, C., Bateman, O. A., Naylor, C. E., Cronin, N., *et al.*, Crystal structures of alpha-crystallin domain dimers of alphaB-crystallin and Hsp20. *J Mol Biol* 2009, *392*, 1242-1252.

[9] Horwitz, J., Alpha-crystallin can function as a molecular chaperone. *Proceedings of the National Academy of Sciences* 1992, *89*, 10449-10453.

[10] Raju, M., Santhoshkumar, P., Krishna Sharma, K., Alpha-crystallin-derived peptides as therapeutic chaperones. *Biochimica et Biophysica Acta (BBA) - General Subjects* 2016, *1860*, 246-251.

[11] Wang, K., Spector, A., alpha-crystallin prevents irreversible protein denaturation and acts cooperatively with other heat-shock proteins to renature the stabilized partially denatured protein in an ATP-dependent manner. *Eur J Biochem* 2000, *267*, 4705-4712.

[12] Zhao, L., Chen, X. J., Zhu, J., Xi, Y. B., *et al.*, Lanosterol reverses protein aggregation in cataracts. *Nature* 2015, *523*, 607-611.

[13] Makley, L. N., McMenimen, K. A., DeVree, B. T., Goldman, J. W., *et al.*, Pharmacological chaperone for alpha-crystallin partially restores transparency in cataract models. *Science* 2015, *350*, 674-677.

[14] Chen, X.-J., Hu, L.-D., Yao, K., Yan, Y.-B., Lanosterol and 25-hydroxycholesterol dissociate crystallin aggregates isolated from cataractous human lens via different mechanisms. *Biochemical and Biophysical Research Communications* 2018, *506*, 868-873.

[15] Shen, X., Zhu, M., Kang, L., Tu, Y., *et al.*, Lanosterol Synthase Pathway Alleviates Lens Opacity in Age-Related Cortical Cataract. *Journal of Ophthalmology* 2018, *2018*, 4125893.

[16] Xu, J., Wang, H., Wang, A., Xu, J., *et al.*, βB2 W151R mutant is prone to degradation, aggregation and exposes the hydrophobic side chains in the fourth Greek Key motif. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 2021, *1867*, 166018.

[17] Kang, H., Yang, Z., Zhou, R., Lanosterol Disrupts Aggregation of Human γD-Crystallin by Binding to the Hydrophobic Dimerization Interface. *Journal of the American Chemical Society* 2018, *140*, 8479-8486.

[18] Yang, X., Chen, X.-J., Yang, Z., Xi, Y.-B., *et al.*, Synthesis, Evaluation, and Structure–Activity Relationship Study of Lanosterol Derivatives To Reverse Mutant-Crystallin-Induced Protein Aggregation. *Journal of Medicinal Chemistry* 2018, *61*, 8693-8706.

[19] Zhou, H., Yang, Z., Tian, X., Chen, L., *et al.*, Lanosterol Disrupts the Aggregation of Amyloid-β Peptides. *ACS Chemical Neuroscience* 2019, *10*, 4051-4060.

[20] Chemerovski-Glikman, M., Mimouni, M., Dagan, Y., Haj, E., *et al.*, Rosmarinic Acid Restores Complete Transparency of Sonicated Human Cataract Ex Vivo and Delays Cataract Formation In Vivo. *Scientific Reports* 2018, *8*, 9341.

[21] Felici, A., Mengato, D., Falciani, M., Bertelli, E., Lanosterol Eye Drops in a Human Juvenile Nuclear Cataract. *Open Journal of Clinical and Medical Case Reports* 2018, *2*, 12–15.

[22] Nagai, N., Fukuoka, Y., Sato, K., Otake, H., *et al.*, The Intravitreal Injection of Lanosterol Nanoparticles Rescues Lens Structure Collapse at an Early Stage in Shumiya Cataract Rats. *International Journal of Molecular Sciences* 2020, *21*, 1048.

[23] Nagai, N., Umachi, K., Otake, H., Oka, M., *et al.*, Ophthalmic In Situ Gelling System Containing Lanosterol Nanoparticles Delays Collapse of Lens Structure in Shumiya Cataract Rats. *Pharmaceutics* 2020, *12*, 629.

[24] Shanmugam, P. M., Barigali, A., Kadaskar, J., Borgohain, S., *et al.*, Effect of lanosterol on human cataract nucleus. *Indian journal of ophthalmology* 2015, *63*, 888-890.

[25] Daszynski, D. M., Santhoshkumar, P., Phadte, A. S., Sharma, K. K., *et al.*, Failure of Oxysterols Such as Lanosterol to Restore Lens Clarity from Cataracts. *Scientific Reports* 2019, *9*, 8459.

[26] Rasmussen, T., van de Weert, M., Jiskoot, W., Kasimova, M. R., Thermal and acid denaturation of bovine lens α-crystallin. *Proteins: Structure, Function, and Bioinformatics* 2011, *79*, 1747-1758.

[27] Surewicz, W. K., Olesen, P. R., On the thermal stability of .alpha.-crystallin: z new insight from infrared spectroscopy. *Biochemistry* 1995, *34*, 9655-9660.

[28] Bloemendal, H., de Jong, W., Jaenicke, R., Lubsen, N. H., *et al.*, Ageing and vision: structure, stability and function of lens crystallins. *Progress in Biophysics and Molecular Biology* 2004, *86*, 407-485.

[29] Steadman, B. L., Trautman, P. A., Lawson, E. Q., Raymond, M. J., *et al.*, A differential scanning calorimetric study of the bovine lens crystallins. *Biochemistry* 1989, *28*, 9653-9658.

[30] Das, B. K., Liang, J. J. N., Chakrabarti, B., Heat-induced conformational change and increased chaperone activity of lens a-crystallin. *Current Eye Research* 1997, *16*, 303-309.

[31] Ghosh, K. S., Pande, A., Pande, J., Binding of γ-Crystallin Substrate Prevents the Binding of Copper and Zinc Ions to the Molecular Chaperone α-Crystallin. *Biochemistry* 2011, *50*, 3279-3281.

[32] Srinivas, P., Narahari, A., Petrash, J. M., Swamy, M. J., Reddy, G. B., Importance of eye lens α-crystallin heteropolymer with 3:1 αA to αB ratio: Stability, aggregation, and modifications. *IUBMB Life* 2010, *62*, 693-702.

[33] Ghahramani, M., Yousefi, R., Khoshaman, K., Moghadam, S. S., Kurganov, B. I., Evaluation of structure, chaperone-like activity and protective ability of peroxynitrite modified human α-Crystallin subunits against copper-mediated ascorbic acid oxidation. *International Journal of Biological Macromolecules* 2016, *87*, 208-221.

[34] Ahmad, M. F., Singh, D., Taiyab, A., Ramakrishna, T., *et al.*, Selective Cu2+ binding, redox silencing, and cytoprotective effects of the small heat shock proteins alphaA- and alphaB-crystallin. *J Mol Biol* 2008, *382*, 812-824.

[35] Singh, D., Raman, B., Ramakrishna, T., Rao, C. M., Mixed oligomer formation between human alphaA-crystallin and its cataract-causing G98R mutant: structural, stability and functional differences. *Journal of molecular biology* 2007, *373*, 1293-1304.

[36] Raju, M., Santhoshkumar, P., Henzl, T. M., Sharma, K. K., Identification and characterization of a copper-binding site in αA-crystallin. *Free Radic Biol Med* 2011, *50*, 1429-1436.

[37] Rasia, R. M., Bertoncini, C. W., Marsh, D., Hoyer, W., *et al.*, Structural characterization of copper(II) binding to alpha-synuclein: Insights into the bioinorganic chemistry of Parkinson's disease. *Proceedings of the National Academy of Sciences of the United States of America* 2005, *102*, 4294-4299.

[38] Bothra, A., Bhattacharyya, A., Mukhopadhyay, C., Bhattacharyya, K., Roy, S., A fluorescence spectroscopic and molecular dynamics study of bis-ANS/protein interaction. *Journal of biomolecular structure & dynamics* 1998, *15*, 959-966.

[39] Li, K., Forciniti, D., Solubility of Lanosterol in Organic Solvents and in Water–Alcohol Mixtures at 101.8 kPa. *Journal of Chemical & Engineering Data* 2020, *65*, 436-445.

[40] Pagano, Rodrigo S., López Medus, M., Gómez, Gabriela E., Couto, Paula M., *et al.*, Protein Fibrillation Lag Times During Kinetic Inhibition. *Biophysical journal* 2014, *107*, 711-720.

[41] Pedersen, J. S., Flink, J. M., Dikov, D., Otzen, D. E., Sulfates dramatically stabilize a salt-dependent type of glucagon fibrils. *Biophysical journal* 2006, *90*, 4181-4194.

[42] Yang, H., Yang, S., Kong, J., Dong, A., Yu, S., Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nat Protoc* 2015, *10*, 382-396.

[43] Kong, J., Yu, S., Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures. *Acta Biochimica et Biophysica Sinica* 2007, *39*, 549-559.

[44] Jackson, M., Mantsch, H. H., The Use and Misuse of FTIR Spectroscopy in the Determination of Protein Structure. *Critical Reviews in Biochemistry and Molecular Biology* 1995, *30*, 95-120.

[45] Micsonai, A., Wien, F., Kernya, L., Lee, Y.-H., *et al.*, Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences* 2015, *112*, E3095-E3103.

[46] Micsonai, A., Wien, F., Bulyáki, É., Kun, J., *et al.*, BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic acids research* 2018, *46*, W315-W322.

[47] Lin, S. Y., Li, M.-J., Ho, C.-J., pH-dependent secondary conformation of bovine lens α-crystallin: ATR infrared spectroscopic study with second-derivative analysis. *Graefes Arch. Clin. Exp. Ophthalmol.* 1999, *237*, 157-160.

[48] Lamba, O. P., Borchman, D., Sinha, S. K., Shah, J., *et al.*, Estimation of the secondary structure and conformation of bovine lens crystallins by infrared spectroscopy: quantitative analysis and resolution by Fourier self-deconvolution and curve fit. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* 1993, *1163*, 113-123.

[49] Horwitz, J., Alpha-crystallin. *Experimental Eye Research* 2003, *76*, 145-153.

[50] Haley, D. A., Horwitz, J., Stewart, P. L., The small heat-shock protein, alphaB-crystallin, has a variable quaternary structure. *J Mol Biol* 1998, *277*, 27-35.

[51] Braun, N., Zacharias, M., Peschek, J., Kastenmüller, A., *et al.*, Multiple molecular architectures of the eye lens chaperone αB-crystallin elucidated by a triple hybrid approach. *Proceedings of the National Academy of Sciences* 2011, *108*, 20491-20496.

[52] Jobling, A. I., Stevens, A., Augusteyn, R. C., Binding of Dexamethasone by α-Crystallin. *Investigative Ophthalmology & Visual Science* 2001, *42*, 1829-1832.

[53] Stevens, A., Augusteyn, R. C., Binding of 1-anilinonaphthalene-8-sulfonic acid to alpha-crystallin. *Eur J Biochem* 1997, *243*, 792-797.

[54] Augusteyn, R. C., Ghiggino, K. P., Putilina, T., Studies on the location of aromatic amino acids in alpha-crystallin. *Biochimica et biophysica acta* 1993, *1162*, 61-71.

SECTION

2. CONCLUSIONS AND FUTURE WORK

2.1. CONCLUSIONS

This dissertation covers a range of topics all related directly or indirectly to protein solubility and its consequences during the processing of proteins or in human health. A considerable effort was put in a supportive study (solubility of sterols) because the data was needed to better understand the effect of sterols in protein solubility. The main conclusions of this work as summarized in the following paragraphs.

In paper I, glycosylated and non-glycosylated mAbs were precipitated by PEG. Effects of PEG molecular weight, pH, and temperature were studied. The glycosylated mAbs had higher solubility than non-glycosylated mAbs in all conditions. Although glycosylation increases the molecular weight of mAbs the glycosylated mAbs required higher amounts of PEG to precipitate, which contradicts the predictions of excluded volume theory. Glycosylation had major effects on the PEG precipitation efficiency but not on the onset precipitation point. The lectin and mAbs were found to form a hetero-complex at several pH levels due to electrostatic attractions. A few solubility models used in this work were not able to explain the effect of glycosylation or protein solubility.

In paper II, the solubility of lanosterol in organic solvents and water-alcohol binary systems was measured at different temperatures. The activity coefficient models of Wilson, NRTL and UNIQUAC correlated the experimental data satisfactorily. The Apelblat –Jouyban-Acree and the van't Hoff–Jouyban-Acree models correlated water-

alcohol binary systems without significant deviations. The lanosterol solubility data collected in this paper showed that the lanosterol concentration used in crystallins aggregates restoration experiments was exceeded the solubility limit of the sterol.

In paper III, it was found that lanosterol and 25-hydroxycholesterol failed to prevent α-crystallin aggregation upon heating. A concentration of 125μM of the two sterols promoted the aggregation of α-crystallin and of the α/γ complex possible by serving as nucleation sites. The chaperone activity of α-crystallin was not affected by the two sterols, which suggest that the sterols do not bind to the chaperone binding site. This was confirmed by the bis-ANS results. FTIR and far-UV CD revealed that the secondary and tertiary structures of α-crystallin were not affected in the presence of the two sterols regardless of concentration. Furthermore, those two sterols had no significant effect on the Cu2+ binding capacity of α-crystallin. The hydrophobic sites were not the interaction sites for those two sterols and the α-crystallin. Thus, no evidence was found that those two sterols interacted with the α-crystallin specifically.

2.2. FUTURE WORK

In this dissertation, several questions were answered but more issues arose as the consequence of this work. The role of glycosylation in protein precipitation by PEG was clarified in the work, but the result was restricted to mAbs. More glycosylated proteins and their non-glycosylated counterparts need to be considered. Then the effect of glycosylation can be added to PEG precipitation models quantitatively. Moreover, pegylation is a popular protein modification that uses PEG instead of glycans to

covalently attach to protein molecules. Then the role of pegylation in the PEG precipitation may be investigated.

In paper II, we collected lanosterol solubility data. As shown in the sterols restoration of cataract patterns, there are more sterols with restoration capabilities. The solubility data of a few sterols in that pattern are still unavailable. The missing solubility data is an obstacle to further study of those sterols. The future work should generate more solubility data of sterols.

Publications reported the failure of lanosterol to restore lenses' transparencies during the course of this dissertation. During the same time, more compounds were discovered to dissolve crystallin aggregates.[43, 44] The mechanism behind this phenomenon was still unclear. The methods used for this dissertation yielded information about the sterols' effects on the crystallin aggregation. Those methods could be extended to other compounds besides lanosterol and 25-hydroxycholesterol.

Senile cataract was used as a target in sterols restoration experiments. Over the lifetime of an individual, crystallins are subjected to a wide range of post-translational modifications that reduced their stabilities and promoted aggregations. [29] Truncation, deamidation, racemisation, oxidation, and methylation are common post-translational modifications that appear in cataract-affected eye lenses. In this dissertation, it showed that the two sterols did not interact with α-crystallin specifically. To better understand the interaction between sterols and senile cataract lenses, the aggregation processes of degraded crystallins should be tested with lanosterol and 25-hydroxycholesterol individually.

Even though the mechanism of crystallin aggregation dissolved by sterols was unclear, it is a reasonable assumption that sterols may dissolve other protein aggregates. Cataract development is not the only disease caused by the protein aggregation. Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and prion diseases are all associated to protein aggregates.[45] The exploratory study done by Zhou *et al.*[46] showed that lanosterol disrupted the aggregation of amyloid-β Peptides, which is associated with Alzheimer's disease. The effects of the sterols were not limited to lanosterol and 25-hydroxycholesterol on those disease-related protein aggregates, and those need to be investigated.

BIBLIOGRAPHY

[1] Moll, R., Divo, M., Langbein, L., The human keratins: biology and pathology. *Histochem Cell Biol* 2008, *129*, 705-733.

[2] Bloemendal, H., de Jong, W., Jaenicke, R., Lubsen, N. H., *et al.*, Ageing and vision: structure, stability and function of lens crystallins. *Progress in Biophysics and Molecular Biology* 2004, *86*, 407-485.

[3] Berg JM, T. J., Stryer L., Section 3.2, Primary Structure: Amino Acids Are Linked by Peptide Bonds to Form Polypeptide Chains., *Biochemistry. 5th edition*, W H Freeman, New York 2002.

[4] Berg JM, T. J., Stryer L. , Section 3.3, Secondary Structure: Polypeptide Chains Can Fold Into Regular Structures Such as the Alpha Helix, the Beta Sheet, and Turns and Loops., *Biochemistry. 5th edition*, W H Freeman, New York 2002.

[5] Berg JM, T. J., Stryer L., Section 3.4, Tertiary Structure: Water-Soluble Proteins Fold Into Compact Structures with Nonpolar Cores., *Biochemistry. 5th edition.*, W H Freeman, New York 2002.

[6] Berg JM, T. J., Stryer L., Section 3.5, Quaternary Structure: Polypeptide Chains Can Assemble Into Multisubunit Structures., *Biochemistry. 5th edition*, W H Freeman, New York 2002.

[7] Gibson, T. J., McCarty, K., McFadyen, I. J., Cash, E., *et al.*, Application of a High-Throughput Screening Procedure with PEG-Induced Precipitation to Compare Relative Protein Solubility During Formulation Development with IgG1 Monoclonal Antibodies. *Journal of pharmaceutical sciences* 2011, *100*, 1009-1021.

[8] Novák, P., Havlíček, V., 4 - Protein Extraction and Precipitation, in: Ciborowski, P., Silberring, J. (Eds.), *Proteomic Profiling and Analytical Chemistry (Second Edition)*, Elsevier, Boston 2016, pp. 51-62.

[9] J.F., Z., *Solubility of Proteins. In: Functionality of Proteins in Food.*, Springer, Berlin, Heidelberg 1997.

[10] Zhang, J., Protein-Protein Interactions in Salt Solutions, *Protein-Protein Interactions - Computational and Experimental Tools*, InTech, Rijeka, Croatia 2012.

[11] Sho Asakura, F. O., Interaction between Particles Suspended in Solutions of Macromolecules. *Journal of polymer Science* 1958, *XXXIII*, 183-192.

[12] Pauly, D., Chacana, P. A., Calzado, E. G., Brembs, B., Schade, R., IgY technology: extraction of chicken antibodies from egg yolk by polyethylene glycol (PEG) precipitation. *J. Vis. Exp.* 2011, 3084-3090.

[13] Boncina, M., Rescic, J., Vlachy, V., Solubility of lysozyme in polyethylene glycol-electrolyte mixtures: the depletion interaction and ion-specific effects. *Biophys. J.* 2008, *95*, 1285-1294.

[14] Zhao, Y., Kang, L., Gao, S., Gao, X., *et al.*, PEG precipitation coupled with chromatography is a new and sufficient method for the purification of botulinum neurotoxin type B [corrected]. *PLoS One* 2012, *7*, e39670-e39670.

[15] Millar, A. H., Heazlewood, J. L., Giglione, C., Holdsworth, M. J., *et al.*, The Scope, Functions, and Dynamics of Posttranslational Protein Modifications. *Annu. Rev. Plant Biol.* 2019, *70*, 119-151.

[16] Fitchette, A.-C., Cabanes-Macheteau, M., Marvin, L., Martin, B., *et al.*, Biosynthesis and Immunolocalization of Lewis a-ContainingN-Glycans in the Plant Cell. *Plant Physiol.* 1999, *121*, 333-344.

[17] Rademacher, T. W., Parekh, R. B., Dwek, R. A., GLYCOBIOLOGY. *Annu. Rev. Biochem.* 1988, *57*, 785-838.

[18] van der Kant, R., Karow-Zwick, A. R., Van Durme, J., Blech, M., *et al.*, Prediction and Reduction of the Aggregation of Monoclonal Antibodies. *J Mol Biol* 2017, *429*, 1244-1261.

[19] Yin, L., Chen, X., Tiwari, A., Vicini, P., Hickling, T. P., The Role of Aggregates of Therapeutic Protein Products in Immunogenicity: An Evaluation by Mathematical Modeling. *J. Immunol. Res.* 2015, *2015*, 401956.

[20] Moussa, E. M., Panchal, J. P., Moorthy, B. S., Blum, J. S., *et al.*, Immunogenicity of Therapeutic Protein Aggregates. *J. Pharm. Sci.* 2016, *105*, 417-430.

[21] Moremen, K. W., Tiemeyer, M., Nairn, A. V., Vertebrate protein glycosylation: diversity, synthesis and function. *Nature Reviews Molecular Cell Biology* 2012, *13*, 448-462.

[22] Atha, D. H., Ingham, K. C., Mechanism of precipitation of proteins by polyethylene glycols. Analysis in terms of excluded volume. *J. Biol. Chem.* 1981, *256*, 12108-12117.

[23] Mahadevan, H., Hall, C. K., Statistical-mechanical model of protein precipitation by nonionic polymer. *AIChE J* 1990, *36*, 1517-1528.

[24] Odijk, T., Depletion Theory and the Precipitation of Protein by Polymer. *J. Phys. Chem. B* 2009, *113*, 3941-3946.

[25] Moreau, K. L., King, J. A., Protein Misfolding and Aggregation in Cataract Disease and Prospects for Prevention. *Trends in molecular medicine* 2012, *18*, 273-282.

[26] Ray, N. J., Biophysical chemistry of the ageing eye lens. *Biophys Rev* 2015, *7*, 353-368.

[27] Truscott, R. J., Friedrich, M. G., The etiology of human age-related cataract. Proteins don't last forever. *Biochimica et biophysica acta* 2016, *1860*, 192-198.

[28] Xiangjia Zhu, A. K., Roger J.W. Truscott, Age-Dependent Denaturation of Enzymes in the Human Lens: A Paradigm for Organismic Aging? *Rejuvenation Research* 2010, *13*, 553-560.

[29] Roodhooft, J. M., Leading causes of blindness worldwide. *Bulletin de la Societe belge d'ophtalmologie* 2002, 19-25.

[30] Zhao, L., Chen, X. J., Zhu, J., Xi, Y. B., *et al.*, Lanosterol reverses protein aggregation in cataracts. *Nature* 2015, *523*, 607-611.

[31] Shanmugam, P., Barigali, A., Kadaskar, J., Borgohain, S., *et al.*, Effect of lanosterol on human cataract nucleus. *Indian Journal of Ophthalmology* 2015, *63*, 888-890.

[32] Makley, L. N., McMenimen, K. A., DeVree, B. T., Goldman, J. W., *et al.*, Pharmacological chaperone for alpha-crystallin partially restores transparency in cataract models. *Science* 2015, *350*, 674-677.

[33] Quinlan, R. A., A new dawn for cataracts. *Science* 2015, *350*, 636.

[34] Shen, X., Zhu, M., Kang, L., Tu, Y., *et al.*, Lanosterol Synthase Pathway Alleviates Lens Opacity in Age-Related Cortical Cataract. *Journal of Ophthalmology* 2018, *2018*, 4125893.

[35] Xu, J., Wang, H., Wang, A., Xu, J., *et al.*, βB2 W151R mutant is prone to degradation, aggregation and exposes the hydrophobic side chains in the fourth Greek Key motif. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* 2021, *1867*, 166018.

[36] Kang, H., Yang, Z., Zhou, R., Lanosterol Disrupts Aggregation of Human γD-Crystallin by Binding to the Hydrophobic Dimerization Interface. *Journal of the American Chemical Society* 2018, *140*, 8479-8486.

[37] Daszynski, D. M., Santhoshkumar, P., Phadte, A. S., Sharma, K. K., *et al.*, Failure of Oxysterols Such as Lanosterol to Restore Lens Clarity from Cataracts. *Scientific Reports* 2019, *9*, 8459.

[38] Nagai, N., Fukuoka, Y., Sato, K., Otake, H., *et al.*, The Intravitreal Injection of Lanosterol Nanoparticles Rescues Lens Structure Collapse at an Early Stage in Shumiya Cataract Rats. *International Journal of Molecular Sciences* 2020, *21*, 1048.

[39] Nagai, N., Umachi, K., Otake, H., Oka, M., *et al.*, Ophthalmic In Situ Gelling System Containing Lanosterol Nanoparticles Delays Collapse of Lens Structure in Shumiya Cataract Rats. *Pharmaceutics* 2020, *12*, 629.

[40] Haslbeck, M., Peschek, J., Buchner, J., Weinkauf, S., Structure and function of α-crystallins: Traversing from in vitro to in vivo. *Biochimica et Biophysica Acta (BBA) - General Subjects* 2016, *1860*, 149-166.

[41] Horwitz, J., Alpha-crystallin can function as a molecular chaperone. *Proceedings of the National Academy of Sciences* 1992, *89*, 10449-10453.

[42] chemical, C. Product Information of Lanosterol, 2016.

[43] Chemerovski-Glikman, M., Mimouni, M., Dagan, Y., Haj, E., *et al.*, Rosmarinic Acid Restores Complete Transparency of Sonicated Human Cataract Ex Vivo and Delays Cataract Formation In Vivo. *Scientific Reports* 2018, *8*, 9341.

[44] Nagai, N., Fukuoka, Y., Sato, K., Otake, H., *et al.*, The Intravitreal Injection of Lanosterol Nanoparticles Rescues Lens Structure Collapse at an Early Stage in Shumiya Cataract Rats. *Int. J. Mol. Sci.* 2020, *21*, 1048.

[45] Ross, C. A., Poirier, M. A., Protein aggregation and neurodegenerative disease. *Nature medicine* 2004, *10 Suppl*, S10-17.

[46] Zhou, H., Yang, Z., Tian, X., Chen, L., *et al.*, Lanosterol Disrupts the Aggregation of Amyloid-β Peptides. *ACS Chemical Neuroscience* 2019, *10*, 4051-4060.

VITA

Ke Li was born in Beijing, People's Republic of China (PRC) in 1989. In 2008, Ke relocated to Guangdong, China to complete his bachelor's degree in chemical engineering and technology in South China University of Technology. Upon securing his bachelor's degree in 2012, Ke decided to expand his knowledge and acquire skills to conduct his own research. He moved to United States in 2012 and got admitted into New Jersey Institute of Technology, Newark NJ, and finished his Master degree in 2014. In 2016, he moved to Missouri University of Science and Technology, Rolla, MO to pursue his PhD. He conducted his doctoral studies in the department of Chemical Engineering under Dr. Daniel Forciniti and obtained his PhD in Chemical Engineering in July 2021.