StAR-related lipid transfer domain protein 5 binds primary bile acids.

Danny Létourneau,1 Aurélien Lorin,1 Andrée Lefebvre, Vincent Frappier, Francis Gaudreault, Rafael Najmanovich, Pierre Lavigne,2 and Jean-Guy LeHoux

Département de Biochimie, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1H 5N4

Abstract  Steroidogenic acute regulatory-related lipid transfer (START) domain proteins are involved in the nonvesicular intracellular transport of lipids and sterols. The STARD1 (STARD1 and STARD3) and STARD4 subfamilies (STARD4–6) have an internal cavity large enough to accommodate sterols. To provide a deeper understanding on the structural biology of this domain, the binding of sterols to STARD5, a member of the STARD4 subfamily, was monitored. The SAR by NMR [1H-15N heteronuclear single-quantum coherence (HSQC)] approach, complemented by circular dichroism (CD) and isothermal titration calorimetry (ITC), was used. Titration of STARD5 with cholic (CA) and chenodeoxycholic acid (CDCA), ligands of the farnesoid X receptor (FXR), leads to drastic perturbation of the 1H-15N HSQC spectrum and the identification of the residues in contact with those ligands. The most perturbed residues in presence of ligands are lining the internal cavity of the protein. Ka values of 1.8·10⁴ M⁻¹ and 6.3·10⁴ M⁻¹ were measured for CA and CDCA, respectively. This is the first report of a START domain protein in complex with a sterol ligand. Our original findings indicate that STARD5 may be involved in the transport of bile acids rather than cholesterol.—Létourneau, D., A. Lorin, A. Lefebvre, V. Frappier, F. Gaudreault, R. Najmanovich, P. Lavigne, and J.-G. LeHoux. StAR-related lipid transfer domain protein 5 binds primary bile acids. J. Lipid Res. 2012. 53: 2677–2689.

Supplementary key words  lipid transport • cholesterol metabolism • bile acids • steroidogenic acute regulatory protein • isothermal titration calorimetry • circular dichroism • NMR spectroscopy

Cholesterol is an essential constituent of mammalian cell membranes; it also serves as a precursor to bile acids, steroid hormones, and vitamin D. Cholesterol homeostasis within the body is controlled through different mechanisms involving its uptake, biosynthesis, transport/trafficking, sorting, storage, secretion, and catabolism to bile acids (1–3). In recent years, the focus of many studies has been on mechanisms involved in inclusion and exclusion of cholesterol from various intracellular organelles. In this respect, specialized nonvesicular lipid transporters of the superfamily of proteins with a steroidogenic acute regulatory (StAR)-related lipid transfer (START) domain were shown to be involved in lipids and cholesterol trafficking between intracellular membranes (4–7).

The START superfamily is defined by the presence of a conserved amino acid sequence of typically 210 amino acids that folds into an α/β helix-grip structure forming a hydrophobic pocket for ligand binding (5, 8–16). This module is conserved throughout the evolution and is involved in the transport of ligands, namely lipids, in mammals. Fifteen mammalian proteins, divided into six subfamilies possess a START domain (4, 5, 11, 17, 18), and two of these subfamilies, STARD1 and STARD4, are reported to bind sterols. The STARD1 subfamily is composed of STARD1 and STARD3. STARD1 (StAR) is the archetype of START domain-containing protein; it binds cholesterol, possesses a mitochondrial leader peptide, and is involved in the transfer of cholesterol into mitochondria in steroidogenic tissues (11, 19–30) and in hepatocytes (31–33).

The second STARD1 subfamily member, STARD3 (metastatic lymph node 64, MLN64), also binds cholesterol (16). STARD3 is a membrane protein that is targeted to the late endosomes by an N-terminal region domain (34). The subcellular localization of STARD1 and STARD3 suggests different roles in cellular cholesterol trafficking between these two proteins. It was also suggested that STARD3 might serve to maintain cholesterol at the membrane of late endosomes prior to its shuttle to cytoplasmic acceptor(s) through the START domain (34).

Abbreviations: CA, cholic acid; CD, circular dichroism; CDCA, chenodeoxycholic acid; CSD, chemical shift displacement; CSI, chemical shift index; ER, endoplasmic reticulum; FXR, farnesoid X receptor; HSQC, heteronuclear single-quantum coherence; ITC, isothermal titration calorimetry; NOE, nuclear Overhauser effect; StAR, steroidogenic acute regulatory; START, steroidogenic acute regulatory-related lipid transfer.

1D. Létourneau and A. Lorin contributed equally to this work.

2To whom correspondence should be addressed.

e-mail: Pierre.Lavigne@USherbrooke.ca

The online version of this article (available at http://www.jlr.org) contains supplementary data in the form of two figures.
The members of STARD4 subfamily, STARD4, STARD5, and STARD6, contain 205–233 amino acid residues and share 26–32% identity between each other (18). In contrast to STARD1 and STARD3, STARD4, STARD5, and STARD6 do not have an N-terminal targeting sequence to direct them to specific cellular organelles; they are thus essentially composed of a START domain. In vitro studies indicated that STARD4 and STARD5 are able to bind cholesterol (6, 35). STARD6 expression has been revealed in the testis (germinal cells); it is not present in steroidogenic tissues (18, 36–38). STARD6 is also widely distributed in the nervous system (36, 39, 40), and it was mainly detected in developing Purkinje cells (41).

In contrast to selective cholesterol binding by STARD1 and STARD3, STARD4 also binds 7α-hydroxycholesterol and 7-hydroperoxycholesterol, whereas STARD5 was reported to also bind 25-hydroxycholesterol (6). STARD4 expression has been revealed in several tissues, including liver, kidneys (18), and keratinocytes (42); STARD5 was also found in macrophages, Kupffer cells, and hepatocytes (43). Recent studies indicate an important role for STARD4 in cholesterol transport and homeostasis. In fact, it was reported that STARD4 overexpression in primary mouse hepatocytes led to a marked increase in intracellular cholesteryl ester concentration by delivering cholesterol to ACAT for esterification (35, 43). Furthermore, STARD4 overexpression enhanced sterol transport to the endocytic recycling compartment and to the endoplasmic reticulum (ER). STARD4 was found very efficient in transporting sterol between membranes in vitro (44). Taken together, these results suggest that cholesterol transport mediated by STARD4 is an important component of the cholesterol homeostasis regulatory machinery.

STARD5 is highly expressed in the liver (Kupffer cells) and kidneys (renal proximal tubules) (6, 18, 45, 46). STARD5 is also present in macrophages, monocytes, promyelocytes, mast cells, and basophils (46). In the human renal proximal tubule cell line HK-2, ER stressors increased STARD5 mRNA levels and induced a relocalization of STARD5 from a diffuse cytoplasmic pattern to a perinuclear and cell periphery distribution (18, 45). STARD5 overexpression promotes an increase in free cholesterol levels in mouse hepatocytes (35). Also, in HK-2 cells, STARD5 expression is higher in cells with greater cholesterol content (17); STARD5 mRNA levels were significantly increased in cholesterol-loaded mouse macrophages (45, 47). The above data support a positive correlation between cellular free cholesterol content and STARD5 expression. Although these studies demonstrate concomitance between cholesterol and STARD5 levels, the actual role of STARD5 remains to be elucidated.

The crystal structures of STARD1 and STARD5 have been recently reported (15); they are almost identical to the crystal structures of the START domains of STARD3 and STARD4 and of those involved in the transfer of ceramides and phospholipids, namely, STARD2, STARD11, STARD13, and STARD14 (9, 10, 13–16). However, to date, no structure of a START domain in complex with a sterol has been determined. In addition, the mechanism of the binding and release of ligand has yet to be fully elucidated.

To improve our understanding of STARD5 structure, dynamics and functions, we undertook the characterization of STARD5 bound to sterol by circular dichroism (CD) and by multinuclear and multidimensional solution-state NMR spectroscopy. Very unexpectedly, we found no indications of cholesterol binding to STARD5. However, we report for the first time that STARD5 can specifically bind cholic acid (CA) and chenodeoxycholic acid (CDCA). As expected from a specific ligand binding reaction, STARD5 has increased thermodynamic stability in presence of CA and CDCA. Moreover, by monitoring the chemical shift displacement (CSD) of backbone amide correlations on 1H-15N heteronuclear single-quantum coherence (HSQC), we were able to locate residues inside the cavity participating to CA and CDCA binding. Titration analyses obtained by Isothermal Titration Calorimetry revealed that STARD5 binds CA and CDCA acid with K_a values of 1.8·10^4 M^-1 and 6.3·10^5 M^-1, respectively.

MATERIALS AND METHODS

Cloning, expression, and purification

The cDNA for the human STARD5 was generously provided by the Structural Genomics Consortium (Karolinska Institutet, Stockholm, Sweden). The construct was modified to remove the N-terminal TEV protease cleavage site and to add a hexahistidine tag at the C terminus. This new construct was sequenced and cloned into the expression vector pET-3a (Novagen). For 15N- or 13C, 15N-double labeling, 15N ammonium chloride (1 g/l), and 13C glucose (3 g/l) (Cambridge Isotopes) as the sole nitrogen and carbon sources were used. E. coli BL21(DE3) was transformed with the plasmid, grown at room temperature (25°C) in M9 medium (100% H_2O or 20% H_2O: 80% D_2O), and induced with 1 mM isopropyl-1-thio-b-D-galactopyranoside (IPTG) when OD600 reached 0.6. After induction, cells were incubated for an additional 18 h at room temperature prior to harvesting by centrifugation. Cells were then resuspended in lysis buffer (3 ml/g of pellets; buffer composition: 50 mM K-Phosphate, 500 mM KCl, 10 mM imidazole, 2 mM TCEP, pH 7.4) with 2 mM TCEP, protease inhibitors (complete Mini) and induced with 1 mM isopropyl-1-thio-b-D-galactopyranoside (IPTG) when OD600 reached 0.6. After induction, cells were incubated for an additional 18 h at room temperature prior to harvesting by centrifugation. Cells were then resuspended in lysis buffer (3 ml/g of pellets; buffer composition: 50 mM K-Phosphate, 500 mM KCl, 10 mM imidazole, 2 mM TCEP, pH 7.4) with 2 mM TCEP, protease inhibitors (complete Mini EDTA-free inhibitors; Roche), and 1 mM PMSF, and frozen at −80°C.

Bacterial pellets were lysed by thawing at 37°C followed by addition of lysosyme (2 mg/ml) and DNase (50 μg/ml). The cell lysate was centrifuged at 19000 g for 30 min, and the supernatant was loaded onto a Ni-NTA column (Qiagen) during 2 h at room temperature. The resin was washed twice with lysis buffer, and the STARD5 recombinant protein was eluted with elution buffer (50 mM K-Phosphate, 500 mM KCl, 250 mM imidazole, 2 mM TCEP, pH 7.4). The buffer was exchanged, and the protein was concentrated using Millipore UltraCel ultracentrifugation filters (10,000 Da MWCO; Amicon, Canada) device into NMR buffer (50 mM K-Phosphate, 50 mM KCl, 2 mM TCEP, pH 7.4) complemented with 10% D_2O and 0.01 mM NaN_3. The final concentrations of the NMR samples were between 0.8 and 1.2 mM. The identity and integrity of the final protein sample was confirmed by SDS-PAGE.

For 15N- or 13C, 15N-double labeling, 15N ammonium chloride (1 g/l), and 13C glucose (3 g/l) (Cambridge Isotopes) as the sole nitrogen and carbon sources were used. E. coli BL21(DE3) was transformed with the plasmid, grown at room temperature (25°C) in M9 medium (100% H_2O or 20% H_2O: 80% D_2O), and induced with 1 mM isopropyl-1-thio-b-D-galactopyranoside (IPTG) when OD600 reached 0.6. After induction, cells were incubated for an additional 18 h at room temperature prior to harvesting by centrifugation. Cells were then resuspended in lysis buffer (3 ml/g of pellets; buffer composition: 50 mM K-Phosphate, 500 mM KCl, 10 mM imidazole, 2 mM TCEP, pH 7.4) with 2 mM TCEP, protease inhibitors (complete Mini EDTA-free inhibitors; Roche), and 1 mM PMSF, and frozen at −80°C.

Bacterial pellets were lysed by thawing at 37°C followed by addition of lysosyme (2 mg/ml) and DNase (50 μg/ml). The cell lysate was centrifuged at 19000 g for 30 min, and the supernatant was loaded onto a Ni-NTA column (Qiagen) during 2 h at room temperature. The resin was washed twice with lysis buffer, and the STARD5 recombinant protein was eluted with elution buffer (50 mM K-Phosphate, 500 mM KCl, 250 mM imidazole, 2 mM TCEP, pH 7.4). The buffer was exchanged, and the protein was concentrated using Millipore UltraCel ultracentrifugation filters (10,000 Da MWCO; Amicon, Canada) device into NMR buffer (50 mM K-Phosphate, 50 mM KCl, 2 mM TCEP, pH 7.4) complemented with 10% D_2O and 0.01 mM NaN_3. The final concentrations of the NMR samples were between 0.8 and 1.2 mM. The identity and integrity of the final protein sample was confirmed by SDS-PAGE.
CD spectropolarimetry

CD measurements were performed on a Jasco J-810 spectropolarimeter equipped with a Pelletier-type thermostat. Routine calibration of the instrument was done with an aqueous solution of d-10-(+)-camphor-sulfonic acid at 290.5 nm. Experiments were performed using quartz cells with a path length of 1.0 mm. For CD spectra and temperature denaturation measurements of STARD5, the protein was dissolved in 10 mM phosphate at pH 7.4, to a final concentration of 10 µM. The protein concentration was determined spectrophotometrically at 280 nm using an extinction coefficient of 30,940 M⁻¹ cm⁻¹. The CD spectra presented are the results of the accumulation of 10 scans at 0.1 nm intervals. Scan speeds and time constants were chosen to allow sufficient response time and achieve favorable signal-to-noise ratios. Temperature-induced denaturation curves were performed in the temperature range from 5°C to 95°C with a rate heating of 1°C/min. The raw mdeg values were transformed in mean residue molar ellipticity (deg·cm²·dmol⁻¹) using the following equation: \[ [\theta]_{222} = \text{CD signal (deg)} \cdot \text{MRW/concentration (g/l)} \cdot l \cdot 10, \]
where MRW is the mean residue weight and l is the path length of the CD cell in cm. The determination of the apparent melting temperature (T°), temperature-dependent enthalpy of unfolding (ΔH°u), and temperature-dependent Gibbs free energy of unfolding (ΔG°u) was performed by the simulation of the temperature denaturation curves with a model describing the equilibrium of a two-state unfolding mechanism and assuming a ΔG°u of 1 kcal·mol⁻¹·K⁻¹ (a typical value for a protein the size of STARD5) as described in Roostae et al. (48). Measurements with the ligand in appropriate buffer were performed at a protein to ligand molar ratio of 1:100, and spectra were taken at a molar ratio of 1:100. In cholesterol studies, a 5 mM stock solution in ethanol was used; the final concentration of ethanol was 0.2%. After 90 min equilibrium time, samples were analyzed by CD. Each spectrum was baseline corrected for buffer and ligand.

NMR spectroscopy

NMR experiments were performed at 298 K on a Varian 600 MHz spectrometer equipped with a Z-axis pulsed-field gradient triple resonance probes. The backbone sequence-specific assignments of ¹H, ¹³C, ³¹P, ²²Na, and ¹⁵N, and side chain ¹³C for the STARD5 were obtained using ¹H-¹⁵N HSQC, HNHA, and standard triple resonance NMR experiments [HNACB, HNCA, HNCO, HN(CO)CA, CBCA(CO)NH, CC(CO)NH, HC(CO)NH, HCC(HC-TOCSY)]. [¹H-¹⁵N] nuclear Overhauser effect (NOE) measurements were done by comparing HSQC spectra with 10 s proton saturation and without proton saturation. The experiments were repeated twice, and the average of the two sets is reported.

NMR data were processed using NMRPipe (49) and analyzed with CCPNmr Analysis (50). The assignments are deposited in the BioMagResBank (http://www.bmrb.wisc.edu/) with accession number 17909. The spectra were referenced as described by Wishart et al. (51). The chemical shift values of ²²Na, ¹³C, and ¹⁵N have been corrected for the deuterium isotopic effect using the values described in Gardner et al. (52).

NMR titration experiments were performed at a protein concentration of 0.2 mM in NMR buffer at pH 7.4. NMR titrations were carried out by acquiring ¹H-¹⁵N HSQC spectra on samples of ¹⁵N-labeled STARD5 with ligand concentration ranging from 0 to 0.4 mM, for a final protein:ligand molar ratio of 1:2. The peak intensities of the disappearing and/or appearing peaks as a function of titration progression were measured. The observed difference (ΔOBS) between the intensities of each point and the initial points have been determined and plotted as function of

$$\Delta_{\text{MAX}} = \left(1 - \frac{\text{FB}}{1 - \text{FB}} \right) \cdot \left(1 + \frac{\text{K}_a}{\text{P}} \right) \cdot \left(1 + \frac{\text{L}}{\text{P}} \right) \cdot \left(1 + \frac{\text{K}_a}{\text{P}} \right) - \left(1 + \frac{\text{L}}{\text{P}} \right)$$

(Eq. 1)

However, because $\Delta_{\text{OBS}}/\Delta_{\text{MAX}}$ monitored from the decrease in peak intensities of the free form gives a descending curve (1 - FB), we fitted those curves with Equation 2 to obtain $K_a$:

$$\Delta_{\text{OBS}} = \left(1 - \frac{\text{FB}}{1 - \text{FB}} \right) \cdot \left(1 + \frac{\text{L}}{\text{P}} \right) \cdot \left(1 + \frac{\text{K}_a}{\text{P}} \right) - \left(1 + \frac{\text{L}}{\text{P}} \right)$$

(Eq. 2)

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) measurements were performed using a VP-ITC (GE Healthcare-MicroCal, Northampton, MA) at 25°C. Protein concentration was chosen to have a c value, defined as the product of the binding constant ($K_a$), the protein concentration ([P]), and the stoichiometric parameter (n) between 10 and 100 and a peak height of more than 0.5 pcal/ sec. The ligand concentrations used were below their respective CMC values. Each experiment consisted of 49 injections into the experimental chamber of 6 µl of the ligand solution at 3 min intervals at a stirring speed of 260 rpm. Heats of dilution were subtracted from the raw titration data before analysis. Experiments were performed in triplicate, and data were fitted by least-squares procedures assuming a one-site binding model using Microcal Origin version 7.0.

Molecular modeling and docking

Coordinates for the missing residues in the α helix of the crystal structure of STARD5 (PDB code: 2R55) were generated by introducing them in the model in an α-helical conformation. Then the potential energy in that region was minimized while restraining the rest of the molecule. A second round of minimization was performed on the entire model. This procedure lead to a model that was superimposable on the initial coordinates with a backbone RMSD of less than 0.2 Å. This was done using the InsightII suite (Accelera, CA). We used the program FlexAID to perform the molecular docking (http://bcb.med.usherbrooke.ca). FlexAID uses a genetic algorithm and allows for side-chain and ligand flexibility and optimizes a surface area complementarity-based scoring function. One thousand poses of CA and CDCA in the binding site of STARD5 were generated. The poses were rescored with the MM-GBSA force field (54). The top (lowest potential energy) 100 poses were analyzed to uncover potential recurrent H-bond patterns between the ligands and the residues of the binding site.

RESULTS

The 3D structure and the flexible regions of the crystal structure of STARD5 are conserved in solution

The assigned ¹H-¹⁵N HSQC of STARD5 recorded at 25°C and pH 7.4 are presented in Fig. IA, B. As expected for a stable tertiary structure, the chemical shift dispersion...
is excellent with little overlap. The assignments of the backbone $^1$H, $^{15}$N, and $^{13}$C-chemical shifts have been reported by us recently (55). The secondary chemical shifts for the Cα, Cβ, C′, and Hα are depicted in Fig. 1C–F, respectively. For Cα and C′, positive and negative secondary chemical shifts (difference between the assigned chemical shift of a given residue and its random coil value) indicate α-helical and β-strand conformation, respectively. This relationship is reversed for Cβ and Hα secondary chemical shifts. Using the chemical shift index (CSI) (56, 57), the sequence specific secondary structure elements of STARD5 have been determined (Fig. 1G, H). The location of the secondary structure elements in STARD5 in solution is identical to that observed in the crystal structure (Fig. 2A). Hence, this confirms that the structure is practically identical in the different milieu. Moreover, we determined the extent of motion of the backbone amides on the ps-ns timescale by measuring the $^1$H-$^{15}$N heteronuclear NOE (Fig. 1G). The $^1$H-$^{15}$N heteronuclear NOE is sensitive to internal motions (in addition to the overall molecular tumbling) of the backbone $^{15}$N-$^1$H vector on the ps-ns timescale. In absence of internal motion, the $^1$H-$^{15}$N heteronuclear NOE is maximal (0.87 at 600 MHz). However, the presence of internal motion lowers the $^1$H-$^{15}$N heteronuclear NOE. Hence, the $^1$H-$^{15}$N heteronuclear NOE values for $^{15}$N-$^1$H vectors in stable secondary structure is expected to be close to the maximum, whereas those in flexible loops and unfolded...
regions are expected to tend toward zero and negative values, respectively. Fig. 2B illustrates the flexible regions of STARD5 in absence of ligand. In this representation, for a large value of the heteronuclear NOE, the diameter of the ribbon is minimal, indicating low mobility of the \(^\text{15}N\text{H}\) vector. Inversely, \(^\text{15}N\text{H}\) vectors with low heteronuclear NOE values are depicted by larger ribbon diameters. As can be observed, the regions of STARD5 with regular secondary structures have low mobility on the ps-ns timescale, whereas loops and unstructured regions have higher mobility. Note the presence of significant movement in the loops connecting \(\alpha_1\) and \(\beta_1\), \(\beta_1\) and \(\beta_2\), \(\alpha_2\) and \(\alpha_3\) (and part of \(\alpha_3\)), \(\beta_5\) and \(\beta_6\) (\(\Omega\)-loop 1), and \(\beta_9/10\) and \(\alpha_4\). Similarly, the atoms in the above regions also have the largest B-factors in the crystal structure of STARD5 (PDB code: 2R55), with no electron density observed for part of the loop between \(\alpha_2\) and \(\alpha_3\) (Fig. 2C). This strongly suggests that these regions are true flexible regions in STARD5 on the ps-ns timescale. However, \(\alpha_1\), \(\beta_1\), and \(\beta_2\) do not participate to the internal cavity assigned to the binding site of STARD5 (Fig. 2B, D).

**STARD5 does not bind cholesterol**

We previously demonstrated the binding of cholesterol to STARD1 using CD (48) and NMR (58); essentially, addition of cholesterol led to the thermodynamic stabilization of the START domain of STARD1 and to drastic changes in the chemical shifts of the \(^\text{1H}\text{15N}\) correlations in its HSQC spectrum. However, under the same experimental conditions, addition of cholesterol to STARD5 did not induce noticeable changes in the far-UV CD spectrum, in the thermodynamic stability (shift in the temperature denaturation curve), or in the \(^\text{1H}\text{15N}\) HSQC (see supplementary data). We did not observe binding of 25-hydroxycholesterol either (data not shown). As described in earlier, STARD5 is expressed in Kupffer cells, circulating macrophages, and renal proximal tubules. These cells are exposed to bile acids and are involved in their transport, reabsorption, and metabolism. In this context, we verified whether STARD5 could bind bile acids; we chose to test cholic acid (CA) and chenodeoxycholic acid (CDCA). The rationale for this choice resides in the fact that CA and CDCA are ligands of the farnesoid X receptor (FXR).
involved in the regulation of bile acid levels (59–61) and that STARD5 and FXR are both expressed in macrophages and the kidneys (18, 46, 62, 63).

**STARD5 binds CA and CDCA**

As shown in Fig. 3A, CA (red) and CDCA (blue) induce only minute changes in the far-UV CD spectrum and more pronouncedly on the thermodynamic stability of STARD5 (Fig. 3B, C). Moreover, the $^1$H-$^1$N HSQC of STARD5 is drastically perturbed by the presence of CA and CDCA (Fig. 4A, B). These results clearly indicate binding and prompted us to further characterize, structurally and thermodynamically, the binding reaction.

**Thermodynamic stability of STARD5 and characterization of the binding of CA and CDCA by CD**

To verify the thermodynamic stability of the tertiary structure of STARD5 with and without ligands, we measured its far-UV CD (Fig. 3A, black dashed line) spectrum and its temperature denaturation (Fig. 3B) monitored by CD (see below). The apparent minimum, starting from 202 nm in the far-UV CD spectrum, indicates a significant mixture of $\alpha$-helical and $\beta$-sheet structure. Spectral analysis, using the program SELCON3 (64), allowed us to simulate (data not shown) the experimental spectrum with contents of $\alpha$-helix, $\beta$-sheet, turns, and disordered structures of 0.18, 0.33, 0.24, and 0.25, respectively. These values are in excellent agreement with those estimated directly from the 3D structure (Fig. 2A).

The simulation of the temperature denaturation of STARD5 with a two-state model (see Materials and Methods), in which the folded state of STARD5 is in equilibrium with its unfolded state, is shown in Fig. 3B (black). Fig. 3C (black) illustrates the corresponding population of the unfolded state ($P_u$) versus temperature. From this simulation, we determined apparent values for $T^\circ$ and $\Delta H^\circ_u(T^\circ)$ of 47.3 ± 0.3°C and 150.1 ± 3.2 kcal·mol$^{-1}$, respectively (Table 1). From these parameters and the Gibbs-Helmholtz equation, $\Delta G^\circ_u(25^\circ C)$ and $\Delta G^\circ_u(37^\circ C)$ values of 9.7 ± 0.4 and 4.6 ± 0.3 kcal·mol$^{-1}$ were calculated. The corresponding populations of the unfolded state ($P_u$) are $8 \cdot 10^{-8}$ and $5 \cdot 10^{-4}$ at 25 and 37°C, respectively. This indicates that STARD5 is stably folded at both temperatures.

Although CA (red) and CDCA (blue) induce slight changes in the far-UV CD spectrum of STARD5 (Fig. 3A), no apparent change in secondary structure content as determined by NMR has been observed (vide infra). However, Fig. 3B, C demonstrate that the apparent $T^\circ$ of STARD5 is significantly increased in presence of CA ($\Delta T^\circ = 5.9^\circ C)$ and CDCA ($\Delta T^\circ = 7.1^\circ C$); this indicates that the stabilization free energy provided by the binding CDCA ($\Delta \Delta G^\circ_u(25^\circ C) = 2.5$ kcal·mol$^{-1}$) is slightly larger than that provided by the binding of CA ($\Delta \Delta G^\circ_u(25^\circ C) = 2.2$ kcal·mol$^{-1}$) (Table 1). Since there is a direct relationship between $\Delta T^\circ$ induced by the binding of ligands and their actual $K_a$ (65), the affinity (or $K_a$) of CDCA for STARD5 is apparently larger than that of CA; this is shown to be the case below.

**Localization of the binding site of CA and CDCA in STARD5 by NMR**

To locate the binding sites of CA and CDCA in STARD5, we monitored the titration of the START domain with
STARD5 binds primary bile acids

For the identification of the residues in contact with the bile acids or that undergo a conformational change upon complex formation. This approach is commonly called SAR by NMR (66). The titration of STARD5 with both molecules causes the displacement of numerous cross-peaks

the two bile acids by NMR. More precisely, we determined the backbone amides of STARD5 that are perturbed by the addition of CA and CDCA using the assigned \( ^1H-^{15}N \) HSQC. Because the 3D structure of STARD5 is known, the identification of the perturbed amides allows

Fig. 4. Binding of primary bile salt to STARD5. Overlay of \( ^1H-^{15}N \) HSQC spectra of STARD5 (0.8 mM in NMR buffer at pH 7.4) alone (black) and in the presence of 2 equivalents of (A) CA (red) or (B) CDCA (blue). Residues perturbed are labeled. (C–F) Backbone chemical shift assignments of STARD5 bound to CA. Secondary chemical shifts of \( ^13C/^{1}H \), \( ^13C/^{2}H \), and \( ^13C/^{3}H \) and CSI for STARD5. \( \alpha \)-helices and \( \beta \)-strands are identified by consecutive CSI values of –1 and 1, respectively. (G) Secondary structure profile from CSI values. (H, I) Weighted CSD map upon binding of CA (H) and CDCA (I) to STARD5. Dashed line represented the mean value of CSD for each ligand. (J, K) Mapping of CSD upon binding of CA (J) and CDCA (K). Worm representation with worm radius proportional to the CSD values and coded in a blue-to-red gradient; regions unaffected (blue) have a thinner backbone worm, and regions perturbed (red) have a thicker backbone worm, whereas intermediate regions are white.
(labeled in Fig. 4A, B) on the $^1$H-$^15$N HSQC of STARD5; indicating that the chemical environment of many backbone amides is perturbed by the presence of the bile acids in the complexes. The addition of CA and CDCA causes the displacement of the same cross-peaks on STARD5 HSQCs, demonstrating that a single subset of residues is perturbed by the presence of both ligands and, hence, suggesting the existence of a common binding site.

Interestingly, free and bound states of STARD5 are in the so-called slow exchange regime: the gradual addition of ligand causes a proportional increase in the intensity of the bound cross-peaks and a concomitant decrease in the intensity of the free form (vide infra). Usually, as a rule of thumb, high- to moderate-affinity complexes ($K_d < 10^4$ M$^{-1}$) are in "slow exchange" and low-affinity complexes ($K_d > 100$ M$^{-1}$) are in "fast exchange" on the NMR chemical shift timescale (67). In the latter case, the actual chemical shifts of the cross-peaks is a population weighted average of those of the free and bound states. To unambiguously identify the backbone amides of all the residues in the HSQCs of the bound form of STARD5, we reassigned the backbone $^1$H, $^13$C, and $^15$N chemical shifts of the bound state STARD5 (BMRB access, 18721) as described previously (55). Fig. 4C–F shows the secondary chemical shifts of the Ca, Cβ, and C′ and CSI, respectively. It is important to note, as indicated by the CSI (Fig. 4F), that the presence of the ligand does not induce changes in the location and extent of the secondary structure elements in STARD5. This suggests, as pointed out by CD (Fig. 3), that the 3D structure of STARD5 in solution is minimally affected by the presence of CA and CDCA.

The assignment of the backbone chemical shifts allows the identification of the backbone amide cross-peaks significantly shifted by the presence of CA and CDCA, respectively. From those assignments, CSD profiles can be established (Fig. 4H, I). The CSD is calculated with the following equation: CSD = ((ΔHf) + (ΔN/SW$^1$/SW$^{15}$N))$^{1/2}$, where SW is the spectral width of both dimensions. Significant displacements are attributed to cross-peaks that have CSD values one SD above the mean (dotted lines in Fig. 4H, I). The patterns of CSD are virtually identical for both the ligands and (regions) most perturbed are located around Arg$^{76}$ (ε3), Thr$^{81}$ (β4), Arg$^{109}$ (β5), Val$^{122}$ (β6), Ser$^{132}$ (β7), Cys$^{158}$ (β9), Thr$^{178}$, and Ser$^{196}$. In Fig. 4J, K, worm representations of the CSDs are displayed on the backbone of STARD5 for CA and CDCA, respectively. The residues with significant CSDs are located on contiguous secondary structure elements and residues that define the internal cavity of STARD5 (Fig. 2D). To the best of our knowledge, this is the first experimental validation of a ligand binding site in a START domain of the STARD1 and STARD4 subfamilies.

### Determination of the apparent binding constant (K_a) of CA and CDCA to STARD5 by NMR

As stated above, the titration experiments monitored by NMR indicate that the free and bound forms of STARD5 interchange in the slow-exchange regime on the NMR timescale. In this regime, the cross-peaks of the residues affected by the presence of the ligand in both states can be observed simultaneously on $^1$H-$^15$N HSQC. The intensities of the cross-peaks corresponding to the free and bound forms are directly proportional to the population of both states (Fig. 5A). Hence, the fraction of bound state (FB = $\Delta_{\text{obs}}/\Delta_{\text{max}}$) can be directly determined as a function of ratio of the concentration of ligand over protein (L/P) when monitoring the appearance of the cross-peaks corresponding to the bound state. From such binding isotherms, an apparent $K_a$ can be determined (see Materials and Methods). Inversely, by monitoring the decrease in the intensity of the cross-peaks of the free form, 1-FB can be determined and used to obtain $K_a$. We selected 10 residues in STARD5 that have both their free and bound cross-peaks well resolved to determine the apparent $K_a$ of CA (Fig. 5B) and CDCA (Fig. 5C). The averages values obtained by NMR ($K_a$) for CA and CDCA are 3.7 ($\pm$ 3.1)·10$^4$ M$^{-1}$ and 7.7 ($\pm$ 4.8)·10$^4$ M$^{-1}$, respectively (Table 2). Although not significantly different, as suggested by the temperature denaturation of the complexes (Fig. 3B, C), STARD5 appears to have a larger affinity for CDCA than for CA.

### Determination of the thermodynamics of binding of CA and CDCA to STARD5 by ITC

To further define the binding parameters, we titrated STARD5 with CA and CDCA using ITC, which allows direct determination of the stoichiometry, $\Delta_H(T)$, $\Delta_S(T)$ and $\Delta_G(T)$, and $K_a(T)$. Fig. 5D, F show the binding isotherms corrected for the heat of dissolution and the nonlinear fits (Fig. 5E, G) to a single binding site model for CA and CDCA, respectively. These results indicate that the binding reaction of CA and CDCA is entropy driven with favorable $\Delta_S$ and unfavorable $\Delta_H$, respectively (Table 2). The corresponding $K_a$ values of 1.8 ($\pm$ 0.2)·10$^4$ M$^{-1}$ and 6.3 ($\pm$ 0.2)·10$^4$ M$^{-1}$, for CA and CDCA, respectively, are in agreement with the $K_a$ values above and confirm the fact that STARD5 has a slightly larger affinity for CDCA than for CA.

### Table 1. Thermodynamic parameters of stability of STARD5

<table>
<thead>
<tr>
<th>$T$ (°C)</th>
<th>$\Delta_H(T)$</th>
<th>$\Delta_S(T)$</th>
<th>$\Delta_G(T)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>apo</td>
<td>47.3 (± 0.3)$^b$</td>
<td>150.1 (± 3.2)</td>
<td>9.7 (± 0.4)</td>
</tr>
<tr>
<td>CA</td>
<td>53.2 (± 0.3)</td>
<td>152.1 (± 3.3)</td>
<td>11.9 (± 0.5)</td>
</tr>
<tr>
<td>CDCA</td>
<td>54.4 (± 0.3)</td>
<td>151.1 (± 3.3)</td>
<td>12.2 (± 0.5)</td>
</tr>
</tbody>
</table>

Thermodynamic parameters of stability of STARD5 were obtained from the simulation of the temperature denaturation apo-STARD5 and STARD5 in the presence of CA and CDCA.

Values are in kcal·mol$^{-1}$.

$^b$ Standard deviation of the fits.
Fig. 5. Determination of the $K_a$ of CA and CDCA for STARD5. (A–C) Determination of the $K_a$ of CA and CDCA for STARD5 using $^1$H-$^1$N HSQC spectra performed at a protein concentration of 0.2 mM in NMR buffer (pH 7.4) with a ligand concentration ranging from 0 to 0.4 mM, for a final protein:ligand molar ratio of 1:2. (A) The free state (black cross-peaks) and bound state of STARD5 are interchanging slowly on the NMR timescale. The cross-peaks of the backbone amide correlation of A71 is shown to have different chemical shifts in the free (red cross-peaks) and the bound (black cross-peaks) states, respectively. In the slow exchange regime, the intensity of the cross-peak is proportional to the population of each states. Plotting $I_{OBS}/I_{MAX}$ (observed/maximal differences in intensity) as a function of the ratio of the total concentration of ligand/protein gives the evolution of the fraction bound (FB) or 1-FB when the intensities of the bound and free states are monitored along the titration, respectively. (B, C) Fitting of the isotherms obtained by monitoring the $\Delta_{OBS}/\Delta_{MAX}$ of A71 upon titration of STARD5 with CA (B) and CDCA (C). (D–G) ITC measurements of STARD5 (protein concentration of 0.4 mM and 0.5 mM for CA and CDCA, respectively). Raw data with (D) CA and (F) CDCA at 25°C. Integrated heat changes, corrected for the heat of dilution, fitted to a single site model for (E) CA and (G) CDCA.
Docking of CA and CDCA inside the STARD5 cavity

As established earlier for STARD3 (16), STARD4 (14), and more recently, STARD1 and STARD5 (15), the internal cavity of START domains is large enough to accommodate a sterol molecule (Fig. 2D). Nevertheless, no structure of STARD1 and STARD4 subfamilies in complex with a sterol ligand has been solved thus far. Hence, to generate possible configuration(s) of the complexes between STARD5 and CA and CDCA, we exploited our CSD data (Fig. 4H–K) and combined them with molecular docking studies using the program FlexAID (see Materials and Methods).

In the selection process, we searched for docking configurations in which CA and CDCA formed H-bonds with residues bearing significant CSDs. Note that the CSDs observed are not necessarily caused by a direct H-bond between the backbone amide and the ligands. In fact, the backbone amides of most of the residues affected by the presence of the ligands are involved in secondary structure H-bonds. Moreover, potential H-bond formation by CA (Fig. 6A) and CDCA (Fig. 6B) is limited to the carboxylate in C24 and carbons bearing an hydroxyl group. The representative poses that optimally fulfill the H-bond potential for CA and CDCA in the binding site of STARD5 are shown in Fig. 6C, D. Interestingly, similar features are observed in both complexes; noteworthy are the H-bonds between the γ1-hydroxyl of the side-chain of Ser132 and the C3-hydroxyl, and between the γ1-hydroxyl of Thr101 and the C7-hydroxyl. Also, in both complexes, the C24 carboxylate receives H-bonds from the γ1-hydroxyls of Thr103 and Thr178, leading to the internal solvation of this ionizable moiety. Also, note the H-bond between the Nη2 of the guanidino group of Arg76 and the C12-hydroxyl of CA (absent in CDCA). Interestingly, this H-bond appears to be facilitated by the formation of an intramolecular salt bridge between Arg76 and Asp80.

<table>
<thead>
<tr>
<th></th>
<th>K_{NMR} \times 10^4 M^{-1}</th>
<th>K_{ITC} \times 10^4 M^{-1}</th>
<th>ΔH \text{kJ mol}^{-1}</th>
<th>TΔS \text{kJ mol}^{-1}</th>
<th>ΔG_{n} \text{kJ mol}^{-1}</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>3.7 (± 3.1)</td>
<td>1.8 (± 2.0)</td>
<td>0.68 (± 0.02)</td>
<td>6.49 (± 0.2)</td>
<td>-5.81 (± 0.90)</td>
<td>0.90 (± 0.02)</td>
</tr>
<tr>
<td>CDCA</td>
<td>7.7 (± 4.9)</td>
<td>6.2 (± 0.2)</td>
<td>0.97 (± 0.02)</td>
<td>7.51 (± 0.2)</td>
<td>-6.54 (± 0.90)</td>
<td>0.80 (± 0.02)</td>
</tr>
</tbody>
</table>

Thermodynamic parameters of the binding of CA and CDCA to STARD5 determined by NMR and ITC at 25°C. Values are in kcal mol\(^{-1}\).

DISCUSSION

To gain insight into the function of STARD5, we characterized its solution structure, dynamics (on the ps-ns timescale), and ligand binding using CD, NMR, and ITC. We found that the α/β helix fold of STARD5 in solution is identical to that of the crystal structure (15). Similar to the crystallographic B-factors, motions of significant amplitudes on the ps-ns timescale were observed in the α2–α3 region (Fig. 2B). Whether this suggests an entry or exit route for the ligands remains to be elucidated by more in-depth relaxation studies, such as relaxation dispersion experiments and molecular dynamic simulations with and without ligands.

Using the same techniques (CD and NMR) previously reported for the demonstration of the binding of cholesterol to STARD1 (48, 58), we were unable to detect cholesterol binding to STARD5 (supplementary Fig. I, A–C).
contrasting with a previous report (6). However, we report for the first time that STARD5 specifically binds the primary bile acids cholic acid (CA) and chenodeoxycholic acid (CDCA). We determined that STARD5 has a larger affinity for CDCA than for CA \( K_d = 6.2 \times 10^{-5} \text{ M} \) and \( K_d = 1.6 \times 10^{-5} \text{ M} \), respectively. For comparison, the affinity of STARD5 for CDCA is in the same order of magnitude reported for the bile acids nuclear receptor (FXR), whereas the affinity of STARD5 for CA is 10 times higher than FXR for CA (68).

The chemical structure of bile acids (Fig. 6A, B) differs from that of other steroids by their nonplanar shape due to their cis A/B ring junction. Bile acids are amphipathic molecules with a concave hydrophilic α-face and a convex hydrophobic β-face. The hydroxyl groups oriented toward the α-face and the carboxylic side-chain confer them their hydrophilic character, while the methyl groups oriented toward the β-face afford them their hydrophobic properties (69). CA possesses three hydroxyl groups in position C3, C7, and C12, two methyl groups (C18 and C19), and an acidic group in C24; the 12-hydroxyl group is absent in CDCA (Fig. 6A, B). Particularly, the binding of both bile acids to STARD5 is associated with positive enthalpy (Fig. 5D–G), indicating that the reaction is entropy driven. The positive \( \Delta H^\circ \) observed for CA and CDCA could be caused by lack of complete compensation of the important dehydration enthalpy associated with the burial of the charged carboxylate and the hydroxyls in CA and CDCA; this dehydration enthalpy can amount to up to 70 kcal·mol\(^{-1}\) for a charged group (70). From an entropic point of view, the burial of the hydrophobic side for both bile acids is expected to be favorable with respect to entropy increase of water.

Using an NMR titration approach and CSD data, we identified the residues of STARD5 perturbed by the presence of the ligands and located the binding site of CA and CDCA. These residues are localized in the internal cavity present in all START domains with known structures (Figs. 2 and 6). This is the first experimental validation of the binding of a sterol molecule inside the internal cavity of the START domain of the STARD1 and STARD4 subfamilies. By combining this information with molecular docking, we generated model complexes of STARD5 with CA and CDCA and selected those that best fulfilled the CSD data. In these complexes, all the hydroxyls of CA and CDCA are involved in H-bonds, and the carboxylate is surrounded by side-chains that can help in its internal solvation.

The present study clearly establishes that STARD5 binds CA and CDCA in vitro; the physiological function of this finding remains to be explored. However, STARD5 is highly expressed in Kupffer cells, peripheral macrophages, and kidney proximal tubule cells. These three cell types are bile acid targets and contain specific receptors or transporters for bile acid uptake and secretion. Hence, we can speculate that STARD5 modulates the activity of putative bile acid-regulated molecules by competing for or delivering CDCA.

In conclusion, this is the first detailed report on the elucidation of a START domain-sterol complex. We also define for the first time that STARD5 specifically binds primary bile acids. These findings will contribute to the understanding of the binding mechanism of a START domain to a sterol and will help to gain more insight into the function of STARD5 in sterol homeostasis.

The authors thank Dr Lari Lehtio (Abo Akademi University, Turku, Finland) for kindly providing us with the STARD5 cDNA.

REFERENCES


