

# SCIENTIFIC REPORTS



OPEN

## STARD6 on steroids: solution structure, multiple timescale backbone dynamics and ligand binding mechanism

Received: 28 April 2016

Accepted: 03 June 2016

Published: 24 June 2016

Danny Létourneau<sup>1,2,3,4</sup>, Mikaël Bédard<sup>1,2,3,4</sup>, Jérôme Cabana<sup>1,2,3,4</sup>, Andrée Lefebvre<sup>1</sup>, Jean-Guy LeHoux<sup>1</sup> & Pierre Lavigne<sup>1,2,3,4</sup>

START domain proteins are conserved  $\alpha/\beta$  helix-grip fold that play a role in the non-vesicular and intracellular transport of lipids and sterols. The mechanism and conformational changes permitting the entry of the ligand into their buried binding sites is not well understood. Moreover, their functions and the identification of cognate ligands is still an active area of research. Here, we report the solution structure of STARD6 and the characterization of its backbone dynamics on multiple time-scales through <sup>15</sup>N spin-relaxation and amide exchange studies. We reveal for the first time the presence of concerted fluctuations in the  $\Omega_1$  loop and the C-terminal helix on the microsecond-millisecond time-scale that allows for the opening of the binding site and ligand entry. We also report that STARD6 binds specifically testosterone. Our work represents a milestone for the study of ligand binding mechanism by other START domains and the elucidation of the biological function of STARD6.

The StAR (Steroidogenic Acute Regulatory) protein is the archetype of a family of proteins having a START (STeroidogenic Acute Regulatory-related lipid Transfer) domain, a module of approximately 210 amino acids that adopts an  $\alpha/\beta$  helix grip fold<sup>1–3</sup>. These proteins are expressed in plants and animals and are preserved through the evolution. In humans, there are fifteen different START domain proteins<sup>4</sup>, which are subdivided into six subfamilies: STARD1 (STARD1/D3), STARD4 (STARD4/D5/D6), STARD2 (STARD2/D7/D10/D11), STARD8 (STARD8/D12/D13), STARD14 (STARD14/D15) and STARD9<sup>4,5</sup>. These proteins bind a variety of lipids and sterols<sup>6</sup> and have various patterns of expression and cellular localization. They are involved in the non-vesicular transport of lipids and sterols, their transfer, their metabolism and cell signalling<sup>4,5,7,8</sup>. Some STARTs have known ligands, but many have yet to have their cognate and functional ligands identified. However, members of the STARD1 and STARD4 subfamilies bind sterol molecules and we are reporting that STARD6 specifically binds C19 and C21 steroids<sup>9</sup>.

Whereas the three-dimensional structures of four of the five START domains of the STARD1 and STARD4 subfamilies; STARD1 and STARD5<sup>10</sup>, STARD3<sup>2</sup> and STARD4<sup>11</sup> have been solved, the structure of STARD6 remains to be determined. Moreover, the mechanism of ligand binding and dissociation is not exactly understood. Compared to classical cases of ligand binding, binding of sterol molecules (and other lipids) to START domains is complicated by the fact that the binding site is completely buried inside the  $\alpha/\beta$  helix grip fold. As reviewed elsewhere<sup>12,13</sup>, different conformational changes have been proposed to give access and allow for the exit of ligands but their time-scales and synchrony remains to be determined. These structural changes involve the opening of the  $\Omega_1$  loop, the C-terminal helix and concerted movements between the  $\Omega_1$  loop and the C-terminal ( $\alpha_4$ ) helix<sup>2</sup> as well as between  $\Omega_1$  loop and  $\alpha_3$  helix<sup>14,15</sup>. The possible existence of an intermediate state with an open conformation involving local and reversible structural changes in the  $\Omega_1$  loop and/or the C-terminal helix to allow for the entry of the ligand has also been suggested<sup>16</sup>. Identification and validation of the exact movements

<sup>1</sup>Département de Biochimie, Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, Sherbrooke, Qc., Canada. <sup>2</sup>Institut de Pharmacologie de Sherbrooke, Université de Sherbrooke, Sherbrooke, Canada. <sup>3</sup>PROTEO; Regroupement Stratégique sur la Fonction, la Structure et l'Ingénierie des Protéines, Université Laval, Québec, Qc., Canada. <sup>4</sup>GRASP; Groupe de Recherche Axé sur la Structure des Protéines, McGill University, Montréal, Qc., Canada. Correspondence and requests for materials should be addressed to J.-G.L. (email: jean-guy.lehoux@sherbrooke.ca) or P.L. (email: pierre.lavigne@usherbrooke.ca)

and their time-scales are of the utmost importance for our understanding of ligand binding by START domains and their functions as transporters of single lipid and sterol molecules.

In the present study, we report the NMR solution structure of apo-STAR6 and the dynamical characterization of its backbone in its free and testosterone bound forms on different time-scales using various  $^{15}\text{N}$  spin-relaxation NMR methods and amide exchange kinetics monitored by HSQC based methods. STAR6 adopts a typical  $\alpha/\beta$  helix grip fold in solution and its backbone undergoes dynamical fluctuations on several time-scales. We observe fluctuations within the  $\alpha_3$  and  $\alpha_4$  helices, as well as in the  $\Omega_1$  loop. Moreover, we demonstrate for the first time that these fluctuations occur on the picosecond-nanosecond (ps-ns), microsecond-millisecond ( $\mu\text{s}$ -ms) and longer time-scales. Unbiased molecular dynamic simulations on the  $\mu\text{s}$  time-scale have sampled structural rearrangements of STAR6 consistent with those fluctuations and suggest that they induce the coalescence of the internal cavity and the surrounding solvent, allowing for the ligand entry. This constitutes the first experimental demonstration of a detailed molecular mechanism for ligand binding by a START domain.

## Material and Methods

**Cloning, expression and purification.** The cDNA for the human STAR6 was provided by Invitrogen (Clone ID 100015037), cloned in pET21b and sequenced. The construct contains a hexahistidine tag at the C-terminus. For  $^{15}\text{N}$ - or  $^{13}\text{C}$ ,  $^{15}\text{N}$ -double labeling,  $^{15}\text{N}$  ammonium chloride (1 g/l) and  $^{13}\text{C}$  glucose (2 g/l) (Cambridge Isotopes) as the sole nitrogen and carbon sources were used. *E. coli* BL21(DE3) were transformed with the plasmid and grown at room temperature (23 °C) in M9 medium (100%  $\text{H}_2\text{O}$  or 20%  $\text{H}_2\text{O}$ : 80%  $\text{D}_2\text{O}$ ) and induced with 1 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) when  $\text{OD}_{600}$  reached 0.6. Transformed *E. coli* BL21(DE3) were gradually selected to grow in  $\text{D}_2\text{O}$ , up to 80%, prior to induction. After induction, cells were incubated for an additional 18 hours 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, 1 mM TCEP, pH7.4) with 1 mM TCEP, protease inhibitors (complete Mini EDTA-free inhibitors from Roche) and 1 mM PMSF and frozen at  $-80^\circ\text{C}$ .

Bacterial pellets were lysed by thawing at  $37^\circ\text{C}$  followed by addition of lysozyme (1 mg/ml) and DNase I (50  $\mu\text{g}/\text{ml}$ ). The cell lysate was then centrifuged at 19000 g for 30 min and the supernatant was loaded onto a Ni-NTA column (Qiagen) during 2 hours at room temperature. The resin was washed twice with lysis buffer and the STAR6 recombinant protein was eluted with elution buffer (50 mM K-Phosphate, 500 mM KCl, 250 mM imidazole, 1 mM TCEP, pH7.4). The protein was then exchanged and concentrated using Millipore UltraCel ultracentrifugation filters (10,000 Da MWCO; Amicon Canada) device into the NMR buffer (50 mM potassium phosphate, 50 mM KCl, 1 mM TCEP, pH 7.4) complemented with 10%  $\text{D}_2\text{O}$  and 0,02%  $\text{NaN}_3$ . The final concentrations of the NMR samples were between 0.8 and 1.1 mM. The purity of the final protein sample was confirmed by SDS-PAGE.

## CD spectropolarimetry and determination ligand binding affinity by thermal shift assay.

Circular dichroism (CD) measurements were performed on a Jasco J-810 spectropolarimeter equipped with a Peltier-type thermostat. Experiments were performed using quartz cells with a path-length of 1.0 mm. For CD spectra and temperature-denaturation measurements of STAR6, the protein was dissolved in 10 mM sodium phosphate buffer at pH 7.4, to a final concentration of  $6\mu\text{M}$ . The protein concentration was determined spectrophotometrically at 280 nm using an extinction coefficient of  $25900\text{ M}^{-1}\text{ cm}^{-1}$ . The CD spectra presented are the results of the accumulation of ten scans at 0.1 nm intervals. Scan speeds and time constants were chosen to allow sufficient response time and achieve favourable signal-to-noise ratios. Temperature-induced denaturation curves were performed in the temperature range from  $10^\circ\text{C}$  to  $90^\circ\text{C}$  with a heating rate of  $1^\circ\text{C}/\text{min}$ . The raw mdeg values were transformed in mean residue molar ellipticity ( $\text{deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$ ) using the following equation:

$$[\theta]_{222} = \frac{\text{CD signal (deg)} \cdot \text{MRW}}{\text{concentration} \left(\frac{\text{g}}{\text{L}}\right) \cdot l \cdot 10} \quad (1)$$

where MRW is the mean residue weight and  $l$  is the path length of the CD cell in cm. The apparent melting temperature ( $T^\circ$ ), apparent standard enthalpy of unfolding at  $T^\circ$ , ( $\Delta H_u^\circ(T^\circ)$ ) and apparent temperature dependent Gibbs free energy of unfolding ( $\Delta G_u^\circ(T^\circ)$ ) were determined by the simulation of the temperature denaturation curves with a model describing the equilibrium for a two-state unfolding mechanism as described in Roostae *et al.*<sup>17</sup>. Measurements with the ligand were performed at protein to ligand molar ratios ranging from 1:0.5 to 1:5 and spectra were taken at a molar ratio of 1:5. For the binding studies, stock solutions were prepared in ethanol at a final concentration of ethanol of 0.4%. After 90 min equilibrium time, samples were analysed by CD. Each spectrum was baseline-corrected for buffer and ligand. The values of the apparent dissociation constant  $K_d(T)$  of STAR6-testosterone complexes were determined by the use of thermal shift analysis as described elsewhere<sup>18</sup>. Briefly, the increase in protein stability (or standard free energy of unfolding) due to ligand binding  $\Delta\Delta G_{u,l}^\circ(T)$  is given by the difference in standard free energy of unfolding in the presence of the ligand ( $\Delta G_{u,l}^\circ(T)$ ) and the standard free energy of unfolding of the free protein  $\Delta G_u^\circ(T)$ . As demonstrated in detail by Layton and Hellinga, assuming that the ligand does not bind to the denatured state, the  $\Delta\Delta G_{u,l}^\circ(T)$  is equivalent to the standard free energy of binding ( $\Delta G_b^\circ(T)$ ). In order to determine  $\Delta\Delta G_{u,l}^\circ(T)$ , the change in standard heat capacity of unfolding  $\Delta C_{p,u,l}^\circ$  must be estimated. This is done by estimating  $T^\circ$  and  $\Delta H_u^\circ(T^\circ)$  for each ligand concentration and by plotting the resultant  $\Delta H_u^\circ(T^\circ)$  as a function of  $T^\circ$ .  $\Delta C_{p,u,l}^\circ$  is assumed to be temperature independent and is given by the slope of the linear curve. Then  $\Delta\Delta G_{u,l}^\circ$  values are calculated at the temperature of the melting of the apo-state ( $T_{\text{apo}}^\circ$ ) for various ligand concentrations and the affinity of the interaction is determined by fitting the canonical equation for the free energy of ligand binding in the stoichiometric regime:

$$\Delta\Delta G_u^{\circ}([L_T], [P_T]) = -RT_{apo}^{\circ} \ln\{1 + [(L_T) - [P_T] - 2K_d] + \sqrt{[(L_T) - [P_T] - 2K_d]^2 - 4[P_T][L_T]/2K_d}\} \quad (2)$$

where  $[L_T]$  and  $[P_T]$  are the total concentration of the ligand and protein, respectively.

**NMR spectroscopy.** All 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  $^1\text{H}$ N,  $^{13}\text{C}\alpha$ ,  $^{13}\text{C}\beta$ ,  $^{15}\text{N}$  and  $^1\text{H}\alpha$  and side-chain  $^{13}\text{C}\beta$  for the STARD6 were obtained using  $^1\text{H}$ - $^{15}\text{N}$  HSQC and the following standard triple-resonance NMR pulse sequences: CBCA(CO)NH, CBCANH, HNC(O), HNCA, HNCACB, HN(CO)CA, CC(CO)NH and HNHA. Side-chains  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts were assigned using HCCONH, HCCH-TOCSY and CCONH pulse sequences also included in the Biopack repertoire.

The chemical shift values of  $^{15}\text{N}$ ,  $^{13}\text{C}\alpha$  and  $^{13}\text{C}\beta$  have been corrected for the deuterium isotopic effect as described elsewhere<sup>19</sup>.  $^{15}\text{N}$ -Edited NOESY (mixing times of 50, 150 ms) and  $\text{C}^{13}$  (aliphatic and aromatic)-Edited NOESY (mixing time of 150 ms) were used to assign NOEs. NMR data were processed using NMRPipe<sup>20</sup> and analyzed with CCPNmr Analysis<sup>21</sup>. Samples containing the ligand were prepared from an ethanolic solution of the steroid to a final ratio of 1: 1 and a final ethanol concentration of 2%.

**$^{15}\text{N}$  Spin relaxation.**  $^{15}\text{N}$ - $R_1$ ,  $R_2$  and  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE experiments were recorded using previously described pulse sequences available in the Biopack repertoire<sup>22</sup>. Delays of 0, 30, 60, 90, 120, 150, 250, 350, 450, 550, 650, 800 and 1000 ms and of 0, 10, 30, 50, 70, 90, 110, 130, 150, 170, 250, 350 and 490 ms were used to obtain  $^{15}\text{N}$ - $R_1$  and  $R_2$  respectively.  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE measurements were done by comparing HSQC spectra with and without a 10-second proton saturation.

**Model free Analysis.** In order to characterize the ps-ns time-scale backbone dynamics and explore the potential occurrence of conformational exchange ( $R_{ex}$ ) on the  $\mu\text{s}$ -ms time-scale, we have analysed the  $^{15}\text{N}$  relaxation data recorded as described above and using ModelFree 4.0<sup>23</sup> in conjunction with Fast ModelFree<sup>24</sup>. Prior to proceeding to the ModelFree analysis, the center of mass of the STARD6 pdb structure was translated to coordinate origin using the program PDBINERTIA and the diffusion tensors for spherical and axially symmetric diffusion models were estimated on the basis of  $^{15}\text{N}$ - $R_2/R_1$  values for rigid amides with the program R2R1\_DIFFUSION (AG Palmer: <http://www.palmer.hs.columbia.edu/software/diffusion.html>).  $S^2$  (squared order parameter of the NH bond vector),  $\tau_e$  (internal correlation time) and  $R_{ex}$  (exchange broadening) were derived from the average of 2  $^{15}\text{N}$ - $R_1$ ,  $R_2$  and  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE datasets (for both the apo- and holo-STARD6) using the common models of the spectral density function (*i.e.*  $S^2$ ,  $S^2\text{-}\tau_e$ ,  $S^2\text{-}R_{ex}$ ,  $S^2\text{-}\tau_e\text{-}R_{ex}$ )<sup>25,26</sup> and two-time-scale internal motion<sup>27</sup>. In both cases, the datasets were fitted best with an axially symmetric rotational diffusion tensor<sup>28</sup>. The axially symmetric rotational diffusion is defined by the anisotropy ratio,  $D_{\text{ratio}} = D_{\text{par}}/D_{\text{perp}}$  where  $D_{\text{par}}$  is the rotational diffusion coefficient parallel to the unique axis of the diffusion tensor and  $D_{\text{perp}}$  is the diffusion coefficient perpendicular to it. In such a model, the local  $\tau_m$  depends on the angle  $\alpha$ , *i.e.* the angle between the NH bond vector and the unique axis of the diffusion tensor. The effective correlation time ( $\tau_{m,\text{eff}}$ ) is given by  $1/6D_{\text{iso}}$ , where  $D_{\text{iso}}$  is equal to  $(2D_{\text{perp}} + D_{\text{par}})/3$ .

**Relaxation Dispersion Data Analysis.** The backbone amide  $^{15}\text{N}$  relaxation dispersion profiles were recorded by measuring their effective transverse relaxation rates,  $R_{2,\text{eff}}$ , with a constant time relaxation delay of 50 ms as a function of the refocusing CPMG field strengths,  $\nu_{\text{cpmg}}$ . The pulse sequence found in the Biopack repertoire which was derived from the work of Palmer *et al.*<sup>29</sup> and Kay *et al.*<sup>30</sup> was used.  $R_{2,\text{eff}}$  with  $\nu_{\text{cpmg}}$  ranging from 80 and 1000 Hz, were measured with a repeat at 480 Hz to estimate experimental error. Spectra were processed with NMRPipe<sup>20</sup> and analysed with CCPNmr<sup>21</sup>. The effective transverse relaxation rates ( $R_{2,\text{eff}}(\nu_{\text{cpmg}})$ ), comprising the transverse relaxation rate in the absence of exchange ( $R_2^0$ ), and the component arising from chemical exchange ( $R_{ex}$ ), were calculated via:

$$R_{2,\text{eff}}(\nu_{\text{cpmg}}) = \frac{1}{\tau_{\text{cpmg}}} \ln \frac{I(\nu_{\text{cpmg}})}{I_0} \quad (3)$$

where  $I(\nu_{\text{cpmg}})$  is the peak intensity at a given CPMG field strength and  $I_0$  is the peak intensity without the CPMG relaxation period. Relaxation dispersion curves were obtained by plotting  $R_{2,\text{eff}}$  as a function of  $\nu_{\text{cpmg}}$ . The data were fitted individually considering no-exchange (Eq. 4a) or using the all timescale approximation described by Ishima and Torchia (Eq. 4b)<sup>31</sup> to identify residues affected by conformational exchange:

$$R_{2,\text{eff}}(\nu_{\text{cpmg}}) = R_2^0 \text{ (no exchange)} \quad (4a)$$

or

$$R_{2,\text{eff}}(\nu_{\text{cpmg}}) = R_2^0 + \frac{(p_a p_b \Delta\omega^2 k_{\text{ex}})}{\left[ k_{\text{ex}}^2 + \sqrt{(p_a^2 \Delta\omega^4 + 144\nu_{\text{cpmg}}^4)} \right]} \text{ (alltimescale)} \quad (4b)$$

where  $R_2^0$  is the exchange rate in absence of a conformational exchange contribution ( $R_{ex}$ ) or  $R_{2,\text{eff}}(\nu_{\text{cpmg}})$  when  $\nu_{\text{cpmg}}$  tends to an infinite value,  $p_a$  and  $p_b$  are the populations of the two interchanging states where  $p_a \gg p_b$  and

$p_b$  represent an intermediate and excited state. The  $\Delta\omega$  parameter is the difference in  $^{15}\text{N}$  chemical shifts in both states and  $k_{\text{ex}}$ , the exchange rate, is given by  $k_A + k_B$ , the two rate constants of the exchange equilibrium.  $R_{\text{ex}}$  is given by:

$$R_{\text{ex}} = R_{2,\text{eff}}(\nu_{\text{cpmg}}) - R_2^0 \quad (5)$$

We then grouped residues for cluster analysis based initial  $k_{\text{ex}}$  values obtained, secondary structure elements and inter-residue secondary or tertiary connectivity. Parameters  $k_{\text{ex}}$  and  $p_b$  were shared parameters for each cluster, and  $R_2^0$  and  $\Delta\omega$  were individually determined. Residues presenting crosspeak overlap were discarded from the relaxation dispersion analysis.

**Data processing and analysis.** NMR data were processed using NMRPipe<sup>20</sup> and analysed with CCPNmr Analysis<sup>21</sup>. The  $^{15}\text{N}$ - $R_1$  and  $R_2$  rate constants and their uncertainties were obtained by fitting the decays of the peak heights to the classical two-parameter single exponential using the least squares fitting program in CCPNmr Analysis. The relaxation dispersion profiles were fit to equation 6 with the program GraphPad Prism 6 (<http://www.graphpad.com/>) for cluster analysis and to determine  $k_{\text{ex}}$ ,  $\Delta\omega$  and  $R_{\text{ex}}$  values and their uncertainties. The chemical shift displacements (CSD) were calculated with the following equation:

$$\text{CSD} = \sqrt{\frac{[(\Delta^1\text{H})^2 + (\Delta^{15}\text{N}/(\text{SW}^1\text{H}/\text{SW}^{15}\text{N}))^2]}{2}} \quad (6)$$

where  $\text{SW}^1\text{H}$  and  $\text{SW}^{15}\text{N}$  are the spectral widths in the  $^1\text{H}$  and  $^{15}\text{N}$  dimensions respectively. Significant displacements are ascribed to cross-peaks that have CSD values of one standard deviation (SD) above the mean.

**Structure calculations.** All NOEs were assigned manually and converted into distance restraints using CCPNmr Analysis. The program DANGLE<sup>32</sup> was used to obtain  $\phi$  and  $\psi$  dihedral angles on the basis of backbone and  $^{13}\text{C}_\beta$  chemical shift values. The CLEANEX-PM pulse-sequence<sup>33</sup> has been used to identify fast and slow exchanging backbone amides. H-Bond restraints have been used only for residues showing slow exchanging amides (not observed on the CLEANEX spectra and presenting high protection factors; see below) and found in regions of regular secondary structures. Structures were calculated using the program ARIA2.2<sup>34</sup> in conjunction with CNS 1.21<sup>35</sup>. Of the 1000 conformers calculated, the 20 presenting the lowest-energy, the lowest number of NOE and dihedral violations were refined in water and submitted to PROCHECK\_NMR<sup>36</sup> to validate the structural quality. The final structure ensemble was deposited in the protein databank (PDB: <http://www.rcsb.org/pdb/>) under the identification code 2MOU.

**Amide proton exchange experiments.** Slow exchange of amide protons was monitored by standard hydrogen–deuterium (H/D) exchange experiments, in which H/D exchange was initiated using the quick dissolution of lyophilized sample techniques at 298 K. A series of  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were acquired after hydration of the lyophilized samples in  $\text{D}_2\text{O}$  at various time intervals up to 4 days. Peak intensities were integrated with CCPNmr Analysis. To obtain  $k_{\text{ex,HX}}$  values, a single exponential function was fitted to the time-dependent decay of signal intensities of non-overlapping cross-peaks:

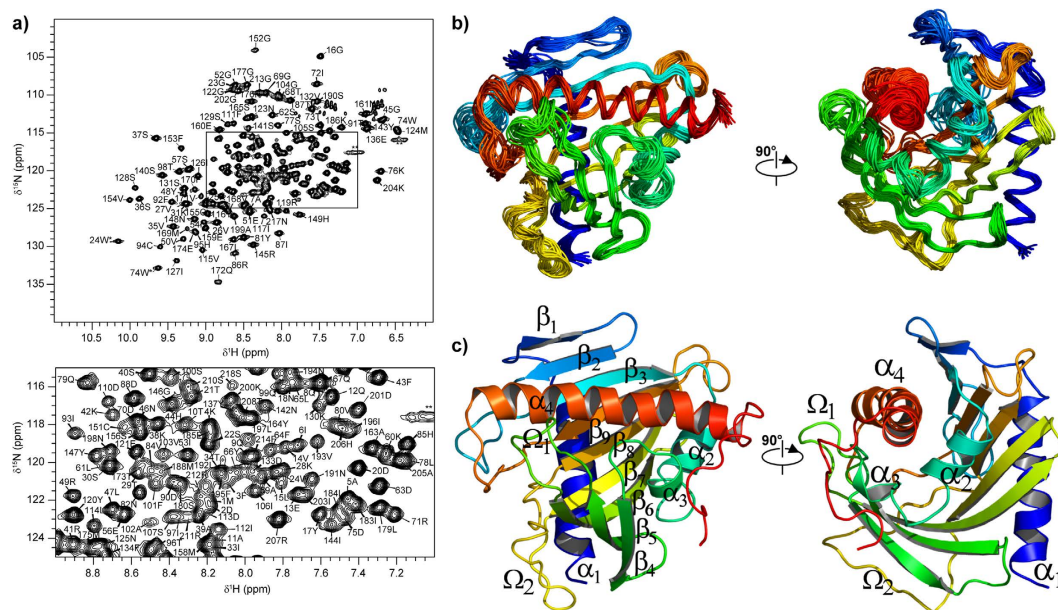
$$I(t) = I(\infty) + A \cdot e^{-k_{\text{ex,HX}}t} \quad (7)$$

where  $I(t)$  is the signal intensity at time  $t$ ,  $I(\infty)$  is the signal intensity at infinite time,  $A$  is the maximum intensity change for a given signal, given by  $I(t)$  at  $t = 0$  minus  $I(\infty)$ , and  $k_{\text{ex,HX}}$  is the site-specific observed rate constant for exchange. Fast exchange of amide protons were identified using the CLEANEX-PM pulse sequence<sup>37</sup>. Each spectrum was recorded with various mixing times (0, 5, 10, 20, 40, 60, 80, 100, 150 and 200 ms). The rates of exchange ( $k_{\text{ex,CLX}}$ ) were determined from fitting the CLEANEX hydrogen exchange build-up curves using Equation 6 with GraphPad Prism 6 (<http://www.graphpad.com/>); where  $\tau_m$  is the mixing times,  $I$  the corresponding peak intensity and  $I_0$  is the reference peak intensity ( $\tau_m = 0$ )<sup>33</sup>.

$$\frac{I}{I_0} = \frac{k_{\text{ex,CLX}}}{R_{1A,\text{app}} + k_{\text{ex,CLX}} + R_{1B,\text{app}}} \left( e^{R_{1B,\text{app}}\tau_m} - e^{(R_{1A,\text{app}}k_{\text{ex,CLX}})\tau_m} \right) \quad (8)$$

The previously determined value of  $0.6 \text{ s}^{-1}$  was used for the apparent relaxation rate of water ( $R_{1B,\text{app}}$ ). Amide proton protection factors ( $P_f$ ) were calculated from the results of all hydrogen exchange experiments as  $k_{\text{int}}/k_{\text{ex,CLX}}$  or  $k_{\text{ex,HX}}$ , where  $k_{\text{int}}$  are the intrinsic exchange rate constants that were estimated using the program SPHERE, based on methods developed by Englander and co-workers (<http://landing.foxchase.org/research/labs/roder/sphere/sphere.html>)<sup>38</sup>.

**Molecular dynamics simulations of apo-STARD6 and its docked complex with testosterone.** The GROMACS software suite<sup>39,40</sup> was used to prepare and run the simulations. The STARD6 protein and its docked complex with testosterone was soaked with the SPC water model<sup>41</sup> and approximately 150 mM NaCl, keeping the net charge of the system at zero. The STARD6 protein was centred in a dodecahedral box in periodic boundary conditions with each side of the box being at least 10 Å from the protein to prevent any interaction between the periodic images of STARD6. The Gromos 54a7 force field was used for the calculations<sup>42,43</sup>. The testosterone molecule and parameters were obtained on the Automated Topology Builder repository (Molecule ID 9924)<sup>44</sup>. The testosterone molecule was docked in the binding site of STARD6 using FlexAID<sup>45</sup> as described in Letourneau *et al.*<sup>46</sup>. Equilibration of the system under conditions of constant number of molecules, volume



**Figure 1.** Solution structure of STARD6 (2MOU). (a)  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectrum of STARD6 (top). Assignments of the cross-peaks are given with the residue number and one-letter code for amino acids. Tryptophan side-chain indole NH cross-peaks are marked with an asterisk (\*). The double asterisks (\*\*) denote protected guanidino ( $\text{H}\eta$ ) protons from arginine side-chains. Detailed view of the more congested central region of the spectrum (bottom). (b) Backbone overlays of the 20 lowest energy structures ensemble. The last 10 residues of the C-terminal were omitted for the sake of clarity. (c) Cartoon representation of the lowest energy structure. Secondary structure elements are labelled. The protein model is rainbow-colored, from the N-terminus in blue to the C-terminus in red.

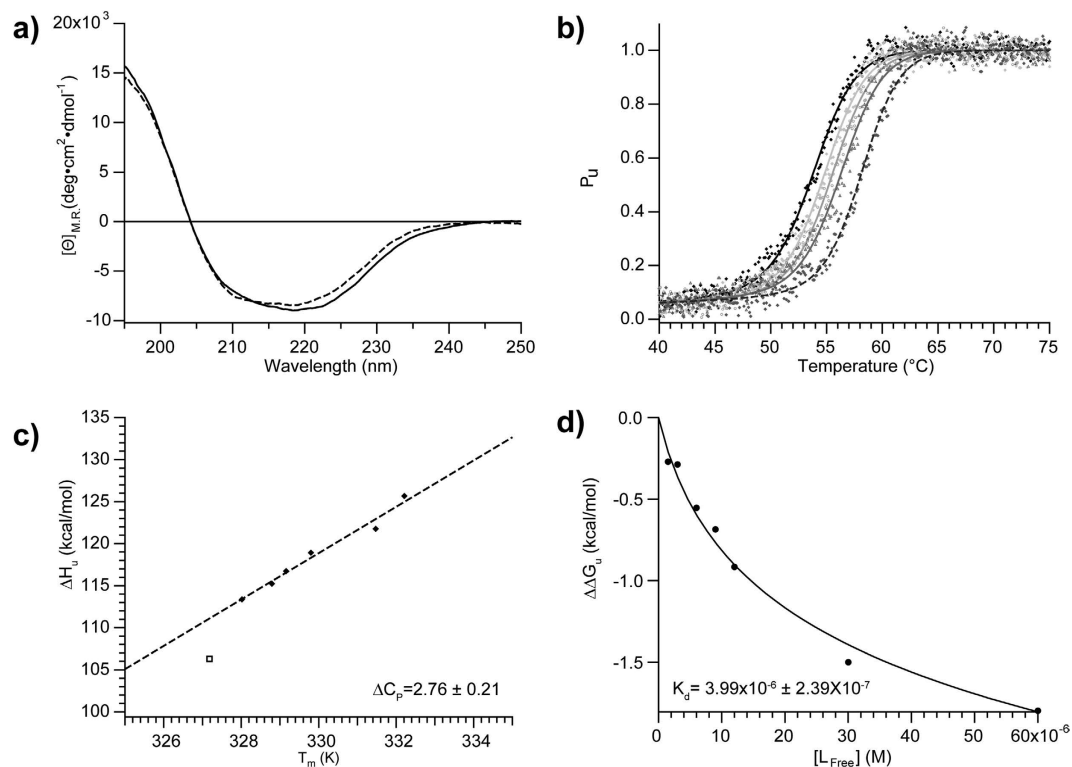
and temperature (NVT) was performed for 10 ps to reach the desired temperature of 310 K. This was followed by equilibration under conditions of constant number of molecules, pressure and temperature (NPT) for 100 ps with the pressure set at 1 bar. Both these steps were performed with position restraints on the heavy atoms of the protein and ligand. A second step of 100 ps of equilibration in NPT condition was performed with position restraint applied only to the backbone atoms of the protein. Unrestrained MD simulations were run in 2 fs steps for 5  $\mu\text{s}$  of total MD simulation time, in the form of 10 simulations of 500 ns in length for each simulated system. The simulations were run in periodic boundary conditions at constant temperature (310 K) and pressure (1 bar) using the Nose-Hoover thermostat<sup>47,48</sup> with  $\tau_T = 0.5$  ps and the Parrinello-Rahman barostat with  $\tau_p = 2$  ps, respectively. Simulated frames were saved every 20 ps, for a total of 250 000 frames per 5  $\mu\text{s}$  simulation. MD trajectories output from GROMACS were converted to PDB files for visual inspection with PyMOL (<https://www.pymol.org/>) and to compressed XTC trajectory files for other analyses. The root mean squared fluctuation (RMSF, a measure of the average backbone mobility over time) analysis was performed using the *g\_rmsf* tool within GROMACS.

## Results and Discussion

**Apo-STARD6 thermodynamical stability.** The  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectrum of STARD6 is well resolved with a single set of resonances (Fig. 1a). This indicates that STARD6 has a stably folded tertiary structure. Accordingly, the thermal denaturation monitored by CD (Fig. 2a,b) fits a two-state model (equilibrium between a folded and unfolded state) and van't Hoff analysis yields a  $T^\circ$  and a  $\Delta H_u^\circ(T^\circ)$  of unfolding of  $54.0 \pm 0.3$  °C and  $106.5 \pm 2.3$  kcal/mol respectively (Table 1). From these parameters and the Gibbs-Helmholtz equation, the free energy change of denaturation ( $\Delta G_u^\circ(T)$ ) at 25 °C and 37 °C are  $9.4 \pm 0.5$  and  $4.6 \pm 0.5$  kcal $\cdot\text{mol}^{-1}$  respectively (Table 1), indicating that the population of native state of STARD6 is virtually 100% at both temperature. Overall, the data presented in this section suggest that STARD6 is stably folded and that if an intermediate state is present, it has a very low population and it is not in slow exchange with the ground state on the chemical shift time-scale.

**The solution structure of apo-STARD6.** The solution structure of apo-STARD6 was determined by NMR spectroscopy using established methods for the assignments of the main chain and side-chain  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shifts and NOE. Amide exchange was used to identify fast and slow exchanging  $\text{H}^{\text{N}}$  atoms and infer their involvement in stable H-Bonds. Overall, 97% of the backbone and 93% of the side-chain atoms ( $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$ ) were assigned leading to 5103 NOE constraints, 302 dihedral angles and 80 H-bonds restraints. The detailed statistics can be found on the BioMagResBank (<http://www.bmrb.wisc.edu/>) with the accession number 18777/19952 and in Table 2.

The 20 lowest-energy structures [PDB ID code 2MOU] are shown superimposed onto the geometric average in Fig. 1b. The structure statistics for the entire protein, including the flexible ends, are summarized in Table 2. The averaged pairwise root-mean-square deviation (RMSD) calculated for backbone-only and for all heavy atoms



**Figure 2.** Far-UV CD spectroscopy of STARD6, experimental determination of  $\Delta C_{p,u,L}^{\circ}$  and apparent  $K_d$  value for STARD6-testosterone complex. (a) Mean residue ellipticity of STARD6 (full line) and in complex with testosterone (dashed line). Each spectrum was baseline-corrected for buffer and ligand. (b) Thermal melting curve recorded at 222 nm and temperature-dependence of the population of unfolded state ( $P_u$ ) for Apo-STARD6 (full line) and in complex at a ratio of 1:2 (light grey), 1:1 (mid grey), 2:1 (dark grey) and 5:1 (dashed line) with testosterone. (c) Estimation  $\Delta C_{p,u,L}^{\circ}$  by a linear fit of  $\Delta H_u^{\circ}(T^{\circ})$  vs.  $T^{\circ}$  values determined experimentally. The value for the apo-STARD6 ( $\square$ ) falls under the fit line, reflecting the absence of the  $\Delta H_b^{\circ}$  contribution to  $\Delta H_u^{\circ}$ . (d) Dependence of  $\Delta\Delta G_u^{\circ}(T)$  on testosterone concentration that can be fit for a  $K_d$  value at the  $T^{\circ}$  of the apo-state using the model in which ligands binds selectively to the native state. The concentration of proteins is  $6\mu\text{M}$  in 10 mM sodium phosphate buffer at pH 7.4. Ligands were dissolved in ethanol were used and the final concentration of ethanol was 0.4%. Samples were analysed after an equilibration period of 90 min.

	$T^{\circ}$ (°C)	$\Delta H_u^{\circ}(T^{\circ})^{a,c}$	$\Delta G_u^{\circ}(25^{\circ}\text{C})^a$	$\Delta G_u^{\circ}(37^{\circ}\text{C})^a$	$K_d^{d,e}$
STARD6 (apo)	54.0 ( $\pm 0.3$ ) <sup>b</sup>	106.5 ( $\pm 2.3$ ) <sup>b</sup>	9.4 ( $\pm 0.5$ )	4.6 ( $\pm 0.5$ )	
Testosterone	58.2 ( $\pm 0.2$ )	117.9 ( $\pm 2.3$ )	16.6 ( $\pm 0.5$ )	9.5 ( $\pm 0.5$ )	3.08 ( $\pm 0.23$ )

**Table 1.** Thermodynamic stability of apo-STARD6, the STARD6-testosterone complex and its apparent dissociation constant obtained from the simulation of the temperature denaturation and thermal shift analysis. <sup>a</sup>Values given in kcal·mol<sup>-1</sup>. <sup>b</sup>Standard deviation of the fits. <sup>c</sup>Calculated at 1:5 ratio. <sup>d</sup>10<sup>-6</sup>M<sup>-1</sup>. <sup>e</sup>at  $T^{\circ}$  of apo-STARD6.

are  $1.18 (\pm 0.18)$  and  $1.8 (\pm 0.14)$  Å, respectively. Those for the START consensus sequence (residues 1–208) are  $0.8 (\pm 0.10)$  and  $1.39 (\pm 0.12)$  Å and those for the regular secondary structure ( $\alpha$  and  $\beta$ ) are  $0.43 (\pm 0.07)$  and  $0.98 (\pm 0.10)$  Å, respectively. Close inspection of Figs 1b and 3h and Supplementary Fig. 1a indicates that the most structurally variable region are located in loops and encompasses residues 38–44, 99–107, 135–142, 174–182 and the C-terminal end of  $\alpha_4$  helix beyond residue 206. Shown in Fig. 1c is a cartoon diagram of the most representative conformer of the ensemble based on its lowest potential energy. Collectively, those results confirm that STARD6 adopt a stable 3D structure consisting of the typical START domain topology.

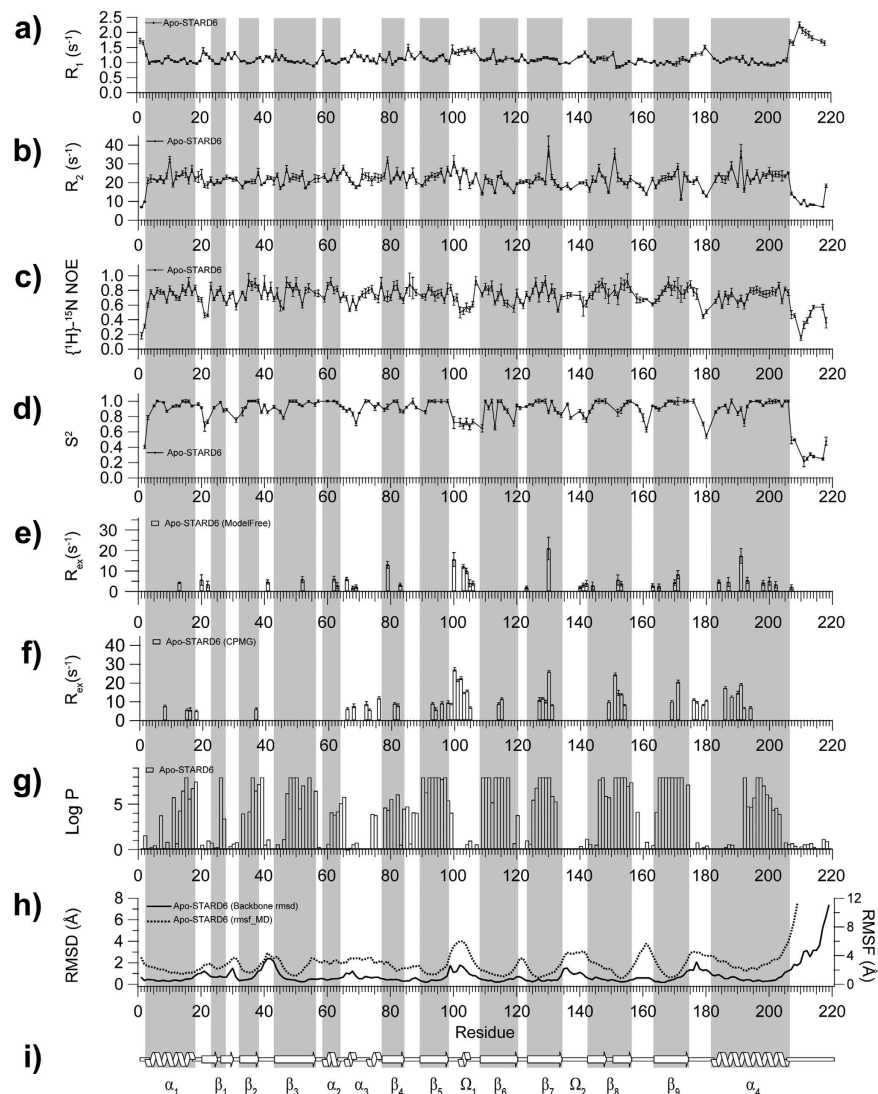
**<sup>15</sup>N-relaxation and picosecond-nanosecond backbone dynamics of apo-STARD6.** <sup>15</sup>N- $R_1$  and  $R_2$  and  $\{^1\text{H}\}$ -<sup>15</sup>N heteronuclear Nuclear Overhauser Enhancement (NOE) were measured at 600 MHz (Fig. 3a–c) and used to characterize the picosecond-nanosecond (ps-ns) dynamics of the backbone of apo-STARD6. The average  $R_1$ ,  $R_2$  and NOE at 600 MHz are  $1.14 (\pm 0.21)$  s<sup>-1</sup>,  $21.3 (\pm 4.7)$  s<sup>-1</sup> and  $0.72 (\pm 0.13)$ , respectively. Higher  $R_1$  and lower NOE and  $R_2$  than the average are observed in the loop between  $\alpha_1$  and  $\beta_1$ , the loop extending to

Restrains for final structure calculations	
NOE distance restraints	
Intraresidue ( $ i-j =0$ )	1612
Sequential ( $ i-j =1$ )	1237
Medium range ( $1 <  i-j  < 5$ )	732
Long range ( $ i-j  \geq 5$ )	1308
Ambiguous NOE	214
Total NOE distance restraints	5103
Hydrogen bonds	$80 \times 2$
$\Phi$ and $\psi$ angles dihedral angle restraints <sup>a</sup>	302
Structure statistics (20 structures)	
Average number of distance violations per structure ( $>0.5 \text{ \AA}$ )	$0.2 \pm 0.4$
Average number of dihedral angle violations per structure ( $>5^\circ$ )	$0.3 \pm 0.4$
RMS deviations from experimental data	
Average distance restraint violation ( $\text{\AA}$ )	$0.043 \pm 0.001$
Average dihedral restraint violation ( $^\circ$ )	$0.60 \pm 0.08$
RMS deviation to mean coordinates	
Backbone heavy atoms ( $\text{\AA}$ )	$1.18 \pm 0.18$
All heavy atoms ( $\text{\AA}$ )	$1.80 \pm 0.14$
For the START domain consensus sequence <sup>b</sup>	
Backbone heavy atoms ( $\text{\AA}$ )	$0.80 \pm 0.10$
All heavy atoms ( $\text{\AA}$ )	$1.39 \pm 0.12$
For secondary structures only	
Backbone heavy atoms ( $\text{\AA}$ )	$0.43 \pm 0.07$
All heavy atoms ( $\text{\AA}$ )	$0.98 \pm 0.10$
Ramachandran plot statistics <sup>c</sup> (%)	
Residues in most favored regions	73.7
Residues in additionally allowed regions	22.8
Residues in generously allowed regions	2.9
Residues in disallowed regions	0.5

**Table 2. Structural statistics for STARD6.** <sup>a</sup> $\Phi$  and  $\psi$  angles were derived from the program DANGLE. <sup>b</sup>START domain consensus sequence comprise residues 1–208. <sup>c</sup>Ramachandran plot statistics were generated using PROCHECK\_NMR for the residues 1–208.

the N-terminal strand of  $\beta_3$ , the short  $\alpha_3$  helix and the beginning of  $\beta_4$ , the  $\Omega_1$  and  $\Omega_2$  loop, the loops between  $\beta_6/\beta_7$  and  $\beta_8/\beta_9$ , the loop before  $\alpha_4$ , as well as the N- and C-termini of the protein (Fig. 3a–c). In accordance with the backbone RMSD (Fig. 3h and Supplementary Fig. 1a), this is indicative of internal motion with significant amplitude on the ps–ns time-scale. On the other hand,  $R_2$  values in the  $\Omega_1$  loop (98–105), structurally juxtaposed to strands  $\beta_7$ ,  $\beta_8$  and the C-terminus of  $\beta_9$ , as well as the N-terminal part of  $\alpha_4$  helix, shows larger values than the average, indicative of conformational exchange on the  $\mu$ s–ms time-scale and/or rotational diffusion anisotropy (Fig. 3b).

To further characterize fast internal motions and conformational exchange on the backbone of STARD6 and to evaluate the existence of rotational diffusion anisotropy, we have analysed the  $^{15}\text{N}$ -relaxation parameters using the model-free approach as described in the methods section. The best statistics for the simulation of the datasets were obtained using an axially symmetric diffusion tensor with a  $D_{\text{ratio}}$  of 1.15 and a  $\tau_{\text{m,eff}}$  of 14.5 ns indicating the existence of sizeable rotational diffusion anisotropy. The  $\tau_{\text{m,eff}}$  estimated is in agreement with the correlation time reported for monomeric protein of similar molecular weight. Accordingly, the agreement between the locations of the  $R_{\text{ex}}$  derived from the model-free analysis and those determined from the relaxation dispersion curves was found to be optimal when the axially symmetric diffusion tensor was used (see below). The structure of STARD6 in the diffusion frame, depicting the unique axis (aligned with the  $\alpha_4$  helix and perpendicular to the  $\alpha_1$  helix) of the diffusion tensor, is presented in Supplementary Fig. 2a. Also shown in Supplementary Fig. 2b, are the derived angles ( $\alpha$ ) between the NH vectors and the unique axis of the diffusion tensor. In absence of conformational exchange ( $R_{\text{ex}} = 0$ ) and internal motions ( $S^2 = 1$ ), NH vectors with  $\alpha$  angle aligned with the unique axis (0 or  $180^\circ$ ) will possess the longest local  $\tau_{\text{m}}$  and highest  $R_2$ . On the opposite, NH vectors with  $\alpha$  angle around  $90^\circ$  will possess the shortest local  $\tau_{\text{m}}$  and lowest  $R_2$ . For a  $D_{\text{iso}}$  of 1.15 and  $\tau_{\text{m,eff}}$  of 14.5 ns, these  $R_2$  values are  $21.2 \text{ s}^{-1}$  and  $19.7 \text{ s}^{-1}$ , respectively. This difference represents the largest effect expected of rotational diffusion anisotropy on  $R_2$  and hence cannot explain the extreme (high) values observed in the  $\Omega_1$  loop, strands  $\beta_7$ ,  $\beta_8$ , the C-terminus of  $\beta_9$  and N-terminal part of  $\alpha_4$  and strongly suggest the existence of conformational exchange in these regions of STARD6.

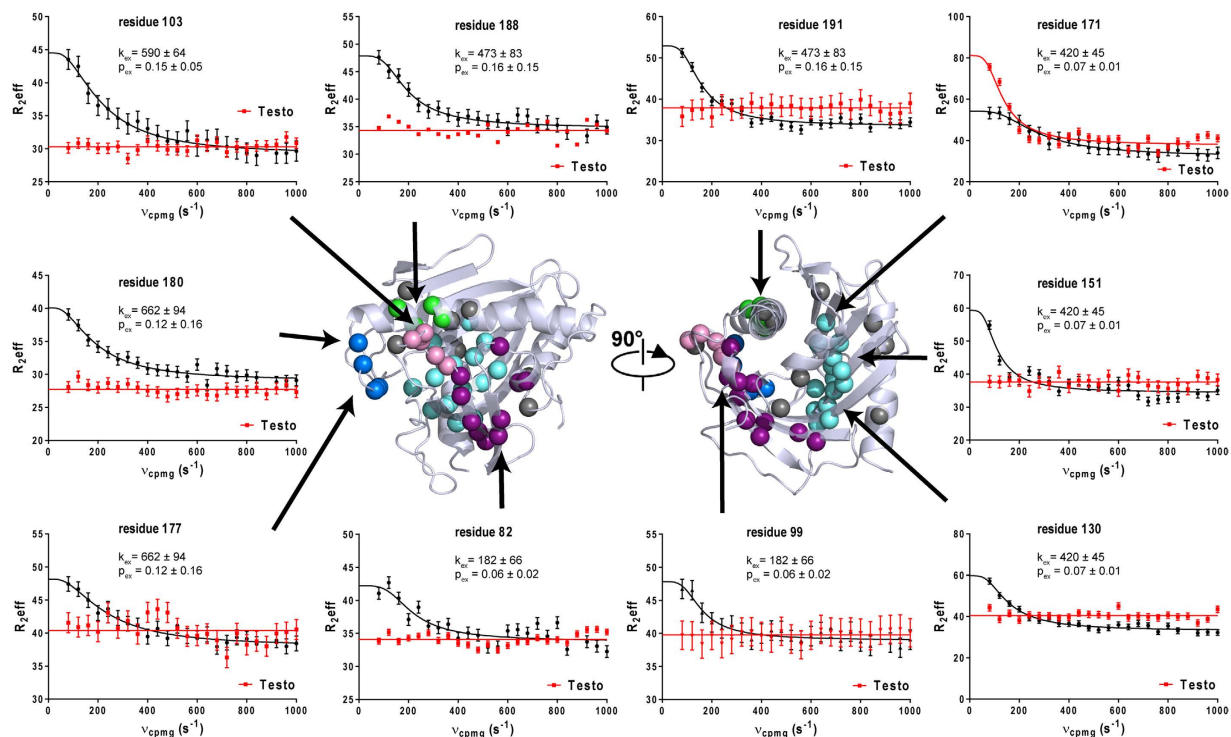


**Figure 3. Backbone dynamics of apo-STARD6.**  $^{15}\text{N}$  relaxation parameters at 600 MHz; (a)  $R_1$ , (b)  $R_2$  and (c)  $\{^1\text{H}\}$ - $^{15}\text{N}$  heteronuclear NOE. (d) Squared order parameter  $S^2$  and (e)  $R_{\text{ex}}$  parameters contribution derived from Model-free analysis. (f) Plot of the  $R_{\text{ex}}$  parameters obtained from the analysis of relaxation dispersion curves. (g)  $\text{H}^{\text{N}}$  protection factors obtained from CLEANEX or standard H/D experiments. (h) Backbone mean RMSD from the 20 lowest energy structures ensemble (black line) and backbone RMSF plot from the MD simulation (dashed line) (i) Secondary structure profile from CSI values for the apo-STARD6.

We report, in Fig. 3d,e, the  $S^2$  and the  $R_{\text{ex}}$  values that best fit the dataset as a function of the residue number according to the model free analysis. As can be observed, the derived  $S^2$  indicate relatively small amplitude motions (average  $S^2$  of  $0.88 \pm 0.10$  for non-overlapping cross peaks) for the backbone across the protein with moderate fluctuation ( $0.8 < S^2 < 0.5$ ) localized in the loops between secondary structure elements (Supplementary Fig. 1b). As could be expected from end fraying effects, more pronounced movements are observed at both terminal ends. Excluding these (and considering only residues 5–207), on average, STARD6 has an  $S^2$  value of 0.91 ( $\pm 0.11$ ), which is significantly higher than the average  $S^2$  ( $0.81 \pm 0.10$ ) of residues in the loops. As one could expect, this indicates that the backbone in the loops is significantly more mobile than those of the core structure on the ps-ns time-scale. More interestingly, the model-free analysis yielded scattered  $R_{\text{ex}}$  with some contiguous values, especially in the  $\Omega_1$  loop and the C-terminal helix where elevated  $R_2$  were observed (Fig. 3e). This suggests that those sections of STARD6 are sensing global motions on the  $\mu\text{s}$ -ms timescale. It is worth mentioning that such motions have been previously shown to be important for the regulation of the activity of different proteins<sup>49–51</sup> and thus, we decided to further characterize them using other approaches.

**$\mu\text{s}$ -ms dynamics of Apo-STARD6.** To confirm the presence and the identity of the backbone amide  $^{15}\text{N}$  undergoing or sensing conformational exchange, we performed a CPMG relaxation dispersion analysis. Representative relaxation dispersion curves are displayed on Fig. 4 and Supplementary Fig. 3. The majority of the residues presenting a  $R_{\text{ex}}$  in the model-free analysis are also found to be affected by conformational exchange





**Figure 4. Microsecond-to-millisecond protein dynamics of apo-STARD6 and in complex with testosterone studied relaxation dispersion at 600 MHz.** Representative dispersion profiles for each cluster of residues displaying  $R_{ex}$  values in the apo-STARD6 form (black curves) and in the presence of ligand (red) are indicated by coloured spheres on the structure. Population ( $P_b$ ) and  $k_{ex}$  values are also displayed. See text for detail. Most residues have their motions dampened except residues 171.

by this technique. Moreover, the analysis of relaxation dispersion curves has led to the detection of a region (176–180) experiencing conformational exchange that was not revealed by the model-free analysis (Fig. 3e,f and Supplementary Fig. 1c). More specifically, these residues are located in the  $\alpha_3$  helix, the C-terminal end of  $\beta_4$  (81–82) and  $\beta_5$  (93–98) strands, the  $\Omega_1$  loop (99–105) and the N-terminal end of  $\alpha_4$  helix (186–194) (Fig. 3f and Supplementary Fig. 1c). In addition, some residues of the  $\beta$ -sheet, for which the side-chains point toward the binding cavity (notably, residues 114–115, 127–131, 149–154 and 169–171), are experiencing conformational exchange (Fig. 3f and Supplementary Fig. 1c). We display on Fig. 4 and Supplementary Fig. 4a,b, the result of a cluster analysis of the dispersion curves. As can be observed, the contiguous residues at the right side of the  $\beta$ -sheet (Fig. 4: cyan spheres), could be fitted with a  $P_b$  of  $0.07 \pm 0.01$  and a  $k_{ex}$  of  $420 \pm 45 \text{ s}^{-1}$ . Within statistical error, we could cluster another set of amides with similar  $k_{ex}$  and  $P_b$  and that belong to residues in the central part of the  $\alpha_4$  helix (green spheres). A cluster with lower apparent  $k_{ex}$  and  $P_b$  involving residues in the  $\alpha_3$  helix and strands at the left edges of the  $\beta$ -sheet was also detected (purple spheres). Comparatively, high B-factors were also observed for the  $\alpha_3$  helix in the crystal structure of other STARD domain<sup>52</sup>. Finally, two other clusters of residues in the  $\Omega_1$  loop (pink spheres) and in the preceding loop of the N-terminal part of the  $\alpha_4$  helix with comparable  $k_{ex}$  ( $590 \pm 64 \text{ s}^{-1}$ ,  $662 \pm 94 \text{ s}^{-1}$ , respectively) and  $P_b$  ( $0.15 \pm 0.05$  and  $0.12 \pm 0.16$ ) values were identified. Residues that we were unable to cluster, based on initial  $k_{ex}$  approximation and location, were not included in the cluster analysis (grey spheres) (Fig. 4 and Supplementary Fig. 4a,b).

**Beyond ms and slower dynamics of the apo-STARD6 form.** In order to complete the characterization of the time-scale and amplitude of the conformational fluctuations present on the backbone of apo-STARD6, we have monitored the  $H^N$  exchange rates of the backbone amides at 25 °C. The rates of exchange ( $k_{ex,CLX}$ ) of fast exchanging amides were measured by the phase-modulated clean chemical exchange pulse scheme. The  $k_{ex,HX}$  of slow exchanging  $H^N$  were measured by the standard hydrogen–deuterium (H/D) exchange experiments, where the exchange is initiated by the quick dissolution of lyophilized sample in  $D_2O$  and monitored by recording  $^1H$ - $^{15}N$  HSQC as a function of time. Protection factors ( $P_f = k_{int}/k_{ex,CLX}$  or  $k_{ex,HX}$ ) were determined as described in materials and methods and are presented on Fig. 3g. Assuming that the exchange rates measured obey to the EX<sub>2</sub> regime, *i.e.* that the closing rate of the opened structure from which the exchange takes places is larger than its intrinsic rate of exchange ( $k_{int}$ ),  $k_{obs}$  depends on the equilibrium constant of the determinant opening reaction ( $K_{op}(T) = e^{-\Delta G_{op}(T)/RT}$ ) and is given by  $P_f^{-1}$ . More than 30% of the total residues have completely exchanged with deuterium within the experimental dead time (30 min). These residues are found to be located in the loops and regions of  $\alpha_3$  and the N-terminal of  $\alpha_4$  as well as on the exposed edges of  $\beta$ -strands (Fig. 3g,i and Supplementary Fig. 1d). The central  $\beta$ -sheet structure and parts of helices  $\alpha_1$  and  $\alpha_4$  contain the most protected amides of the protein. The volumes of those cross-peaks have not significantly decreased even after four days. These amides

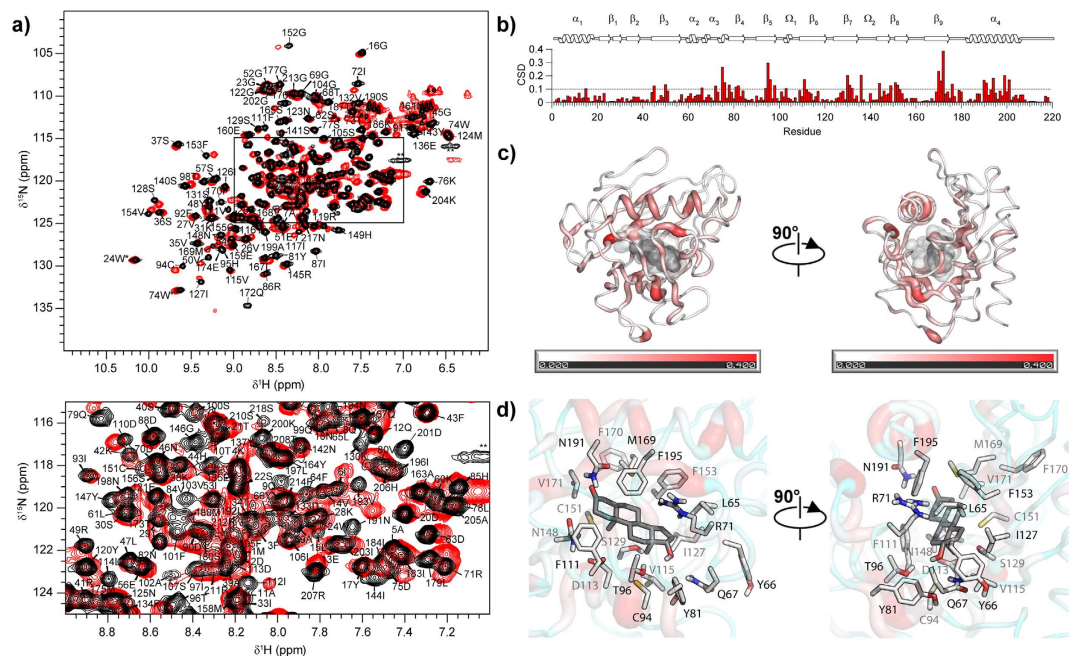
have  $\log P_f$  values over 8. This corresponds to  $\Delta G_{op}$  values equal or larger than 10.5 kcal/mol and a population ( $P_{op}(T) = e^{-\Delta G_{op}(T)/RT}/(1 + e^{-\Delta G_{op}(T)/RT})$ ) of the opened state equal to or smaller than  $10^{-8}$ . This value is in agreement with the  $\Delta G_{op}^{\circ}(25^{\circ}\text{C})$  determined from the temperature denaturation curve for two-state unfolding. As seen on Supplementary Fig. 1d, these residues constitute a central and ring-shaped hydrophobic core around the cavity. Intermediate values of  $\log P_f$  between 3.5–7.5 are observed for contiguous regions. Of interest, helices  $\alpha_2$  and  $\alpha_3$ , strand  $\beta_4$  and the C-terminal part of helix  $\alpha_4$ , which make tertiary contacts, have  $\log P_f$  values near four. Such a protection factor indicates that this portion of the tertiary structure of STARD6 has a smaller and independent  $\Delta G_{op}$  of around 5.4 kcal/mol and a population of the opened state of  $10^{-4}$ . Similar local  $\Delta G_{op}$  for amides at the interface of the  $\alpha_4$  helix and strands  $\beta_1$  and  $\beta_2$  were also observed. As could be expected, amides exposed to the solvent and not potentially involved in H-bonds or more remote from the central core have the fastest exchange rates. At best, they have protection factors with  $\log$  values around 0.5. Such values correspond to  $\Delta G_{op}$  close to 1 kcal/mol and a population of the opened state that can reach up to 0.5. Particularly, one can notice that such weakly protected amides are located in the N-terminal part of the  $\alpha_4$  helix and the  $\Omega_1$  loop (Fig. 3g–i and Supplementary Fig. 1d).

**5  $\mu\text{s}$  Molecular Dynamics Simulation of Apo-STARD6.** In an effort to further explore and rationalize the structural fluctuations detected on all time-scales, we have run a molecular dynamics simulation as described in details in material and methods. Briefly, the energy minimized average structure was soaked in a box of water and its cavity was filled with water molecules. This system was equilibrated for 210 ps before the simulation was started. We present on Fig. 3h, the RMSF calculated for the ensemble of frames collected over a total simulation time of 5  $\mu\text{s}$ . One can notice a general agreement between the RMSD and the RMSF with the lowest RMSF values observed in the central core residues and the highest observed for residues distant from the central core. However, compared to the RMSD, the larger RMSF values indicate the presence of fluctuations larger than those observed on the ns-ps time-scale. In accordance, inspection of the trajectories revealed that, in accordance with its low  $P_f$  and concomitant local fluctuations in the tertiary contacts involving the  $\Omega_1$ , that the N-terminal part of the  $\alpha_4$  helix undergoes local denaturation. Collectively, these large amplitude motions lead to the coalescence of the internal water molecules (and cavity) with the surrounding solvent, suggesting a role for these global motions in the regulation of ligand entry in the binding cavity.

**Binding of testosterone to STARD6.** In a recent analysis of the volumes and the shapes of the cavities and ligands of the START domains from the STARD4 subfamily, we had noticed that the cavity of STARD6 was significantly smaller, suggesting that smaller sterols such as steroids might be potential ligands for STARD6<sup>9</sup>. We demonstrate below that STARD6 can bind sterols with high affinity using testosterone as an example. First, we have used a circular dichroism-based thermal shift assays (see materials and methods) to determine the affinity of testosterone for STARD6 (Fig. 2 and Table 1). As for STARD5<sup>46,53</sup>, no noticeable changes in the spectrum and in secondary structure content are observed for STARD6 in the presence of ligand. However, the thermodynamical stability of STARD6 is significantly increased in the presence of increasing amounts of testosterone as expected for favourable binding of a ligand (Fig. 2b). The quantitative analysis of a two-state equilibrium model from thermal denaturation linked to binding (see materials and methods and Table 1) provided a  $K_d$  values of  $3.08 \times 10^{-6}$  at 54 °C (Fig. 2c,d and Table 1).

**The binding site of STARD6.** The binding site of STARD6 was defined by the classical SAR by NMR approach<sup>54</sup>. All the chemical shifts of the backbone of STARD6 were reassigned upon testosterone binding (BMRB accession number 26763) and the chemical shifts displacements (CSD) of all the backbone amide cross-peaks caused by the presence of the ligand were measured. Numerous cross-peaks on the  $^1\text{H}$ - $^{15}\text{N}$  HSQC of STARD6 are perturbed in presence of testosterone, indicating that their chemical environment is modified following ligand binding (Fig. 5a). As observed for the binding of Cholesterol to STARD1<sup>12,13</sup> and bile acids to STARD5<sup>46,53</sup>, the binding of testosterone to STARD6 is in the slow exchange regime (not shown). In agreement with CD measurements (Fig. 2a), apart from helix  $\alpha_3$  and the  $\Omega_1$  loop, testosterone did not alter significantly the secondary structure of apo-STARD6 (Fig. 2a, Fig. 6i and Supplementary Fig. 5). The secondary chemical shifts for the C $\alpha$ , C $\beta$ , C' and H $\alpha$  as well as the CSI relative to the primary and secondary structures for the apo-STARD6 and STARD6-testosterone complex are shown in Supplementary Fig. 5. From the inspection of the CSDs (Fig. 5b) and the worm representation (Fig. 5c), the most perturbed amides are those that belong to the residues surrounding the cavity or that are next to it (i.e. 72, 75–76, ( $\alpha_3$ ), 78, 81 ( $\beta_4$ ), 94–96, 98 ( $\beta_5$ ), 111 ( $\beta_6$ ), 130–132 ( $\beta_7$ ), 151–153 ( $\beta_8$ ), 170–172 ( $\beta_9$ ), 190–191, 194 and 199 ( $\alpha_4$ )); clearly demonstrating that the testosterone is located inside the cavity. Some remote residues (i.e. 45, 50, 82, 103, 136, 144, 149, 176 and 201) have their chemical shifts also perturbed by the presence of ligand, most likely originating from differential interactions with residues perturbed by the direct interaction with the ligand.

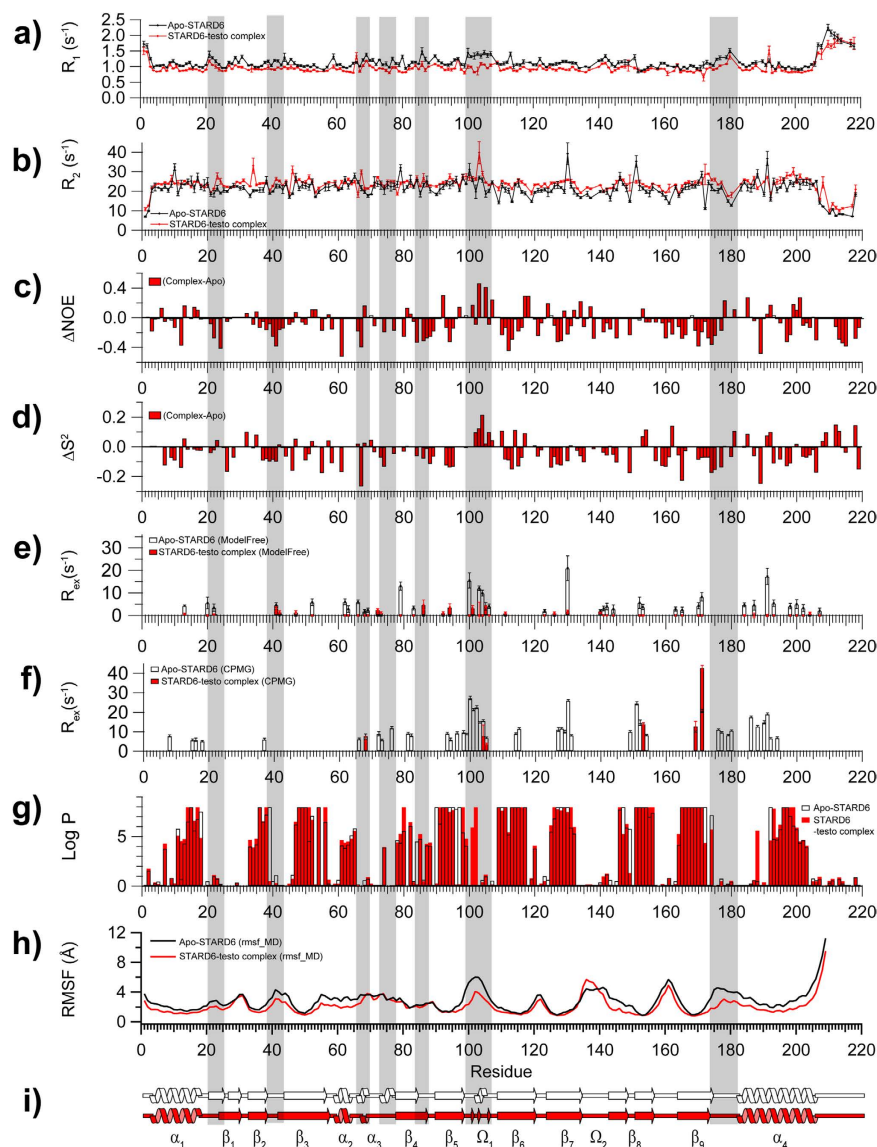
Based on CSD and on the shape of the cavity we have generated a complex between STARD6 and testosterone using molecular docking (Fig. 5c). Inspection of the complex reveals that the central part of testosterone is involved in non-polar interactions involving (C94, F111, V115, I127, C151, F153, M169, F170, V171 and F195) (Fig. 5d). In addition, testosterone possesses two polar groups capable of accepting and donating hydrogen bonds: the carbonyl (C=O) at C3 and a hydroxyl (OH) at C17. Interestingly, residues near both these polar moieties are polar in nature (Q67, R71, S129, N148 and N191) and can engage interaction to fulfill H-Bond potential. Altogether, these observations support the possibility that this configuration of the complex could be stable and highly representative.



**Figure 5. The Binding site of STARD6.** (a) (top) Overlain of  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectra of apo-STARD6 (black) and in the presence of testosterone (red). Detailed view of the more congested central region of the spectra (bottom). (b) Weighted chemical shift displacements (CSD) observed upon testosterone binding to STARD6. Dashed line represented the mean value of CSDs. Significant displacements are assigned to cross-peaks that have CSD values one standard deviation above the mean (dotted lines). (c) Normalized CSDs mapped onto STARD6 structure for testosterone. Worm representation with worm radius proportional to CSD values and coded in a cyan-to-red gradient; regions without a significant perturbation (cyan) have a thinner backbone worm, regions of higher perturbations (red) have a thicker backbone worm, and intermediate regions are white. (d) Molecular docking of the STARD6-testosterone complex in accordance with the CSD data.

**Changes in the backbone dynamics of STARD6 in complex with testosterone.**  $^{15}\text{N}$  relaxation parameters of testosterone-bound STARD6 were best fitted with a  $D_{\text{ratio}}$  of 1.15 and a  $\tau_{\text{m,eff}}$  of 18.5 ns. This is a significant change compared to the  $\tau_{\text{m,eff}}$  of 14.5 ns of the apo-form. Such a reduction in molecular tumbling is manifested by the general decrease in  $R_1$  and increase in  $R_2$  (Fig. 6a,b). Such an effect can be caused by one or a combination of the following phenomena: i) a conformational changes, ii) aggregation or iii) increase of the solvent viscosity. Since no significant change in the secondary chemical shifts of the backbone atoms between the apo- and holo-states have been observed, the possibility of a conformational change appears unlikely. Moreover, the CSD (*vide supra*) observed upon the addition of testosterone are localized around the binding site and do not define any other contiguous surface that could define a specific oligomerization interface. Hence, we trust that this increase in  $\tau_{\text{m,eff}}$  is a result of either non-specific and weak aggregation and/or an increase in viscosity due to a slight excess of testosterone.

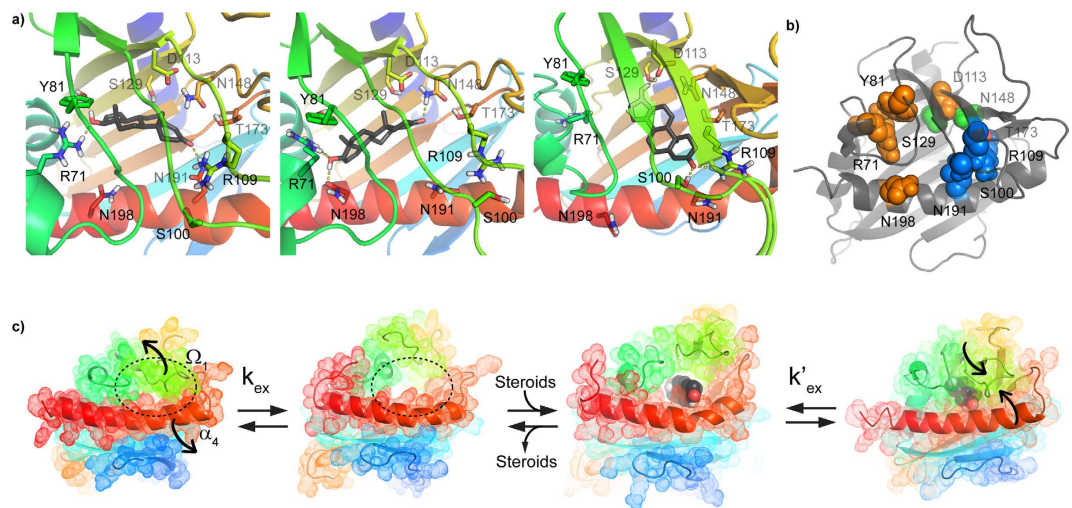
More interestingly, we observed a general decrease of  $S^2$  and  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE in the backbone of the bound state of STARD6 compared to its apo-form. The average values are 0.88 ( $\pm 0.16$ ) (apo) vs. 0.82 ( $\pm 0.15$ ) (holo) for  $S^2$  and 0.72 ( $\pm 0.13$ ) (apo) vs. 0.65 ( $\pm 0.21$ ) (holo) for  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE. More specifically, these gains of flexibility are located in the  $\beta_2/\beta_3$  loop (36–41), the  $\alpha_3$  helix (66–78), the  $\beta_4$  strand (83–88), the  $\beta_5$  strand (90–95), the  $\beta_6$  strand (111–116), the  $\beta_7$  strand (125–132), the  $\Omega_2$  loop (135–142), the end of  $\beta_9$  and into the following loop (150–163), the C-term of  $\beta_9$  (170–177) and the  $\alpha_4$  helix (189–206). However, an increase of rigidity on this time-scale occurs in the  $\Omega_1$  loop region upon ligand binding as denoted by the difference (testosterone-apo) of squared order parameter  $S^2$  and  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE values (Fig. 6c,d and Supplementary Fig. 7a,b). Thus, beside the  $\Omega_1$  loop, the presence of testosterone imparts an increase in the amplitude of fluctuation of the amide bonds on the picosecond-nanosecond time-scale. Interestingly, this is the time-scale associated with the conformational entropy of the backbone of proteins with decreases in  $S^2$  scaling with an increase in conformational entropy<sup>55</sup>. Classically, the formation of a complex is expected to lead to a reduction in conformational entropy and an increase in  $S^2$ . However, as discussed elsewhere, similar decreases in  $S^2$  (hence, increases in conformational entropy) upon complex formation have been reported before and proposed to contribute significantly to the stability of the complex by decreasing its Gibbs free energy by optimisation of its conformational entropy<sup>56</sup>. Accordingly, Fig. 7a displays three different and representative snapshots from the molecular dynamics simulation of the STARD6-testosterone complex in which the ligand adopts different configurations while maintaining non-polar interactions and multiple alternative H-Bonds. The distribution and the occurrence of H-bonds between STARD6 and testosterone are summarized in Fig. 7b and Supplementary Table 1. The main polar interaction between the testosterone hydroxyl group and STARD6 occurs via R71, Y81, D113 and S129; in parallel, the carbonyl group interacts mainly through S100, N191 and S129 (Fig. 7a,b and Supplementary Table 1). Altogether,



**Figure 6. Backbone dynamics changes upon testosterone binding to STARD6.**  $^{15}\text{N}$  relaxation parameters measured at 600 MHz; (a)  $R_1$  and (b)  $R_2$  and (c)  $\{^1\text{H}\}$ - $^{15}\text{N}$  heteronuclear NOE difference (testosterone-apo). (d) Difference (testosterone-apo) in squared order parameters  $S^2$ . (e) Plots of the  $R_{\text{ex}}$  parameters determined from the Model-free analyses for apo-STARD6 (open bar) and STARD6-testosterone complexes (red bar). (f) Plot of  $R_{\text{ex}}$  parameters determined from the analysis of the relaxation dispersion profiles (open bar) and STARD6-testosterone complexes (red bar). (g)  $\text{H}^{\text{N}}$  exchange protection factors from CLEANEX or standard H/D experiments for apo-STARD6 (open bar) and STARD6-testosterone complexes (red bar). (h) Backbone RMSF plot obtained from the MD simulation for apo-STARD6 (black line) and STARD6-testosterone complex (red line) (i) Secondary structure profile from CSI values for apo-STARD6 (white) and STARD6-testosterone complexes (red).

these results suggest that the complex is more dynamical than the apo-form on the ps-ns timescale and that this dynamic behavior is caused by a multitude of potential configurations and intermolecular interactions between testosterone and the side-chains projected inside the STARD6 cavity with seemingly similar probability of existence or Gibbs free energy.

Overlaid to those changes in motions on the ps-ns time-scale, the binding of testosterone quenches larger scale motions in the backbone related to changes in the  $^{15}\text{N}$  chemical shifts and to amide exchange. Indeed, with the exception of a few residues (T68, G104, S105, F153, M169 and V171),  $\mu\text{s}$ -ms motions measured by relaxation dispersion (CPMG) are strongly damped by the presence of the ligand (Figs 4 and 6f and Supplementary Fig. 7d). The model-free analysis corroborates that the testosterone bound form appears less mobile on this time-scale (Fig. 6e). In particular,  $\mu\text{s}$ -ms motions in  $\Omega_1$  loop and the N-terminal side of the  $\alpha_4$  helix are completely abolished in the presence of testosterone (Fig. 6e,f and Supplementary Fig. 7d). Accordingly, the RMSF values obtained from the molecular dynamics simulations of the apo and testosterone-bound form shows a local rigidification of



**Figure 7. Binding mechanism and molecular dynamics simulation.** (a) Representative snapshots taken from trajectories of the molecular dynamics simulation of STARD6-testosterone complex in which the ligand position changes while maintaining H-Bonds and non-polar interactions. Carbon atoms of the STARD6 model are rainbow-colored, from the N-terminus in blue to the C-terminus in red. Carbon atoms of testosterone are colored black. Oxygen, nitrogen and hydrogen atoms are respectively red, blue and white. The backbone is shown in cartoon representation and side chains are shown as sticks. H-bonds are shown as yellow dash. (b) The main polar interactions between the testosterone hydroxyl group and STARD6 occurs via R71, Y81, D113, N198 (orange) and S129 (green), whereas the testosterone carbonyl group interacts mainly through S100, R109, T173, N191 (blue) and N148 (green) or S129. (c) STARD6 is in equilibrium ( $k_{ex}$ ) with an excited state in which fluctuations (cartoon arrows) of the  $\Omega_1$  loop and the N-terminal portion of the  $\alpha_4$  helix induces the coalescence of the internal cavity of apo-STAR6 with the external solvent. Dashed line defines the internal cavity of apo-STAR6 fusing with the solvent. Upon steroids binding, a new equilibrium ( $k'_{ex}$ ) with a more frequent closed state appears. The backbone is shown in cartoon representation and semi-transparent spheres represent the molecular surface of the protein. Testosterone is shown in opaque spheres representation. Colors are the same as in (a).

the  $\Omega_1$  loop upon ligand binding (Fig. 6h and Supplementary Fig. 7e). This is further confirmed by the appearance of a resistance to the amide exchange of residues in  $\Omega_1$  loop (F101 and A102) and the N-terminal of  $\alpha_4$  helix (M188) (Fig. 6g, Supplementary Figs 6 and 7c). Moreover, such high protection factors indicates that the backbone amide of these residues must be involved in the formation of hydrogen bonds or become highly protected by their local environment in the complex. Interestingly, few residues have persistent conformational exchange at the level of their backbone.

In summary, the binding of testosterone to STARD6 causes an overall increase in flexibility on the ps-ns timescale while dampening motions of larger amplitude on the  $\mu$ s-ms timescale. As for the the major urinary protein, this redistribution of the conformational entropy from one timescale to another may contribute to the stabilisation of the complex by the so-called entropy-entropy compensation effect<sup>57,58</sup>. The persistence of fast-motions in the complex may also be important for the re-opening of the cavity and the release of the ligand. Indeed, such a ps-ns to  $\mu$ s-ms motion interplay have already been documented in the literature<sup>49–51,56–58</sup>.

## Conclusion

With the completion of the STARD6 structure presented here, the structure of all the members of the mammalian sterol binding START domains family (STARD1 and STARD4 subfamilies) is now completed. Except for some loop deletions and insertions, the structures of all these START domains are very similar; the backbone superposition of STARD1<sup>10</sup>, STARD3<sup>2</sup>, STARD4<sup>11</sup>, STARD5<sup>10</sup> and STARD6 is shown in Supplementary Fig. 8. Their backbone superimposes onto each other with pairwise RMSDs of less than 2.0 Å (Supplementary Table 2).

We also unveil the time-scales and amplitude of the structural fluctuations proposed to be important for the binding of ligands by START domains (Fig. 7c). Most noteworthy, we demonstrate the concomitant conformational exchange in the  $\Omega_1$  loop and the N-terminal portion of  $\alpha_4$  in the apo-state of STARD6. We show that this leads to the coalescence of the buried binding site with the surrounding bulk water molecules to allow for ligand entry. Despite their low primary structure homologies<sup>12</sup>, we propose that this conformational exchange and binding mechanism, presented here, will be conserved across the fifteen human START domains. In this regard, significant fluctuations (high B-factors) in the  $\Omega_1$  loop are conserved in the structure of the apo-forms of START domains in the majority of the crystal structures reported<sup>2,10,11,14,59</sup>. In addition, steered molecular dynamics simulations have been used to show that the opening of the  $\Omega_1$  loop can play the role of a lid and allow for the exit of cholesterol in STARD1 and STARD3<sup>15</sup>. Also in accordance with the mechanism we propose, Baker *et al.* have engineered a disulphide bridges that links the  $\Omega_1$  loop to the N-terminus of the  $\alpha_4$  helix of STARD1. They have shown that this mutant was only functional under reducing conditions<sup>60</sup>. Finally, we have unveiled the presence of slightly less frequent fluctuations on the  $\mu$ s-ms clustered around the  $\alpha_3$  helix and strands  $\beta_4$  and  $\beta_5$ , lining the

opposite end of the cavity (Fig. 3f, Supplementary Fig. 1c and Fig. 4a,b). High B-factors were also reported in this region of STARD11. It was suggested that these fluctuations participate to the binding and exit of ceramides<sup>14,52</sup>. Alternately, we propose that these fluctuations could be important for the expulsion of water molecules with the concomitant entry of the ligand in the binding site.

Lastly, we reported that STARD6 binds to pregnenolone, precursor of all steroid-hormones<sup>9</sup>. Here, we show that it also binds to testosterone with an affinity in the 10<sup>-6</sup> M range. Together, those results represent an important step towards the understanding of the biological functions of STARD6 and strongly suggest a role for STARD6 in steroid transport. Accordingly, many different *in vivo* studies realized on STARD6 point towards such a function for this protein. Indeed, STARD6 is differentially expressed in germ and brain cells during development and ageing and its nuclear localization suggests that STARD6 might play a regulatory role in cells where it is expressed<sup>61,62</sup>.

In neonatal rats brain, STARD6 is found in Purkinje cells during the first and second week after birth<sup>63–68</sup>, period that Purkinje cells are the most active in steroid synthesis<sup>69</sup>, when cerebellar neural circuit formation occurs<sup>70</sup>. In aged rats, STARD6 slightly increased in the hippocampus but diminished in the cerebrum and cerebellum cortex, suggesting a role for STARD6 during brain ageing<sup>64</sup>.

In adult rat testes, STARD6 is present in the seminiferous tubules and localizes in germ cell nuclei. This START protein is expressed only during the late stages of spermatogenesis, suggesting that STARD6 may play a role during cell maturation<sup>61,62</sup>. In addition, its expression in testes decreases with ageing and may be associated with decreased spermatogenesis in aged rats<sup>61</sup>. By analogy, in the pig ovary, STARD6 is differentially expressed during the estrous cycle, in the mid-luteal phase corpora lutea<sup>71</sup>.

Hopefully, the detailed structural and dynamical data obtained for steroid binding by STARD6 in this study will serve as landmarks for studying others START domain binding mechanism. Beyond the structural biology, the discovery that STARD6 binds steroids with high affinity brings new insight for studying its role in Purkinje and germ cells during development and ageing, and could eventually contribute to the resolution of certain types of pathologies occurring in the brain and those related to fertility problems.

## References

- Iyer, L. M., Koonin, E. V. & Aravind, L. Adaptations of the helix-grip fold for ligand binding and catalysis in the START domain superfamily. *Proteins* **43**, 134–144 (2001).
- Tsujishita, Y. & Hurley, J. H. Structure and lipid transport mechanism of a StAR-related domain. *Nat Struct Biol* **7**, 408–414 (2000).
- Ponting, C. P. & Aravind, L. START: a lipid-binding domain in StAR, HD-ZIP and signalling proteins. *Trends Biochem Sci* **24**, 130–132 (1999).
- Alpy, F. & Tomasetto, C. Give lipids a START: the StAR-related lipid transfer (START) domain in mammals. *J Cell Sci* **118**, 2791–2801 (2005).
- Clark, B. J. The mammalian START domain protein family in lipid transport in health and disease. *J Endocrinol* **212**, 257–275 (2012).
- Schrick, K., Nguyen, D., Karlowski, W. M. & Mayer, K. F. START lipid/sterol-binding domains are amplified in plants and are predominantly associated with homeodomain transcription factors. *Genome Biol* **5**, R41 (2004).
- Alpy, F., Legueux, F., Bianchetti, L. & Tomasetto, C. [START domain-containing proteins: a review of their role in lipid transport and exchange]. *Med Sci (Paris)* **25**, 181–191 (2009).
- Clark, B. J. & Stocco, D. M. In *START Domain Protein Structure and Ligand Specificity* (eds Létourneau, D., Lavigne, P., Lefebvre, A. & Lehoux, J. G.) 49–72 (Springer, 2014).
- Létourneau, D., Lefebvre, A., Lavigne, P. & Lehoux, J. G. The binding site specificity of STARD4 subfamily: Breaking the cholesterol paradigm. *Mol Cell Endocrinol* (2014).
- Thorsell, A. G. *et al.* Comparative structural analysis of lipid binding START domains. *PLoS One* **6**, e19521 (2011).
- Romanowski, M. J., Soccio, R. E., Breslow, J. L. & Burley, S. K. Crystal structure of the *Mus musculus* cholesterol-regulated START protein 4 (StarD4) containing a STAR-related lipid transfer domain. *Proc Natl Acad Sci USA* **99**, 6949–6954 (2002).
- Létourneau, D., Lavigne, P., Lefebvre, A. & Lehoux, J.-G. In *Cholesterol Transporters of the START Domain Protein Family in Health and Disease* (eds Barbara J., Clark & Douglas M., Stocco) Ch. 3, 49–72 (Springer New York, 2014).
- Létourneau, D., Lefebvre, A., Lavigne, P. & Lehoux, J. G. STARD5 specific ligand binding: comparison with STARD1 and STARD4 subfamilies. *Mol Cell Endocrinol* **371**, 20–25 (2013).
- Kudo, N. *et al.* Crystal structures of the CERT START domain with inhibitors provide insights into the mechanism of ceramide transfer. *J Mol Biol* **396**, 245–251 (2010).
- Murcia, M., Faraldo-Gomez, J. D., Maxfield, F. R. & Roux, B. Modeling the structure of the StART domains of MLN64 and StAR proteins in complex with cholesterol. *J Lipid Res* **47**, 2614–2630 (2006).
- Mathieu, A. P., Fleury, A., Ducharme, L., Lavigne, P. & Lehoux, J. G. Insights into steroidogenic acute regulatory protein (StAR)-dependent cholesterol transfer in mitochondria: evidence from molecular modeling and structure-based thermodynamics supporting the existence of partially unfolded states of StAR. *J Mol Endocrinol* **29**, 327–345 (2002).
- Roostaeae, A., Barbar, E., Lehoux, J. G. & Lavigne, P. Cholesterol binding is a prerequisite for the activity of the steroidogenic acute regulatory protein (StAR). *Biochem J* **412**, 553–562 (2008).
- Layton, C. J. & Hellinga, H. W. Thermodynamic analysis of ligand-induced changes in protein thermal unfolding applied to high-throughput determination of ligand affinities with extrinsic fluorescent dyes. *Biochemistry* **49**, 10831–10841 (2010).
- Gardner, K. H., Rosen, M. K. & Kay, L. E. Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR. *Biochemistry* **36**, 1389–1401 (1997).
- Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293 (1995).
- Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
- Farrow, N. A. *et al.* Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry* **33**, 5984–6003 (1994).
- Palmer, A. G., Rance, M. & Wright, P. E. Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance carbon-13 heteronuclear NMR spectroscopy. *Journal of the American Chemical Society* **113**, 4371–4380 (1991).
- Cole, R. & Loria, J. P. FAST-Modelfree: a program for rapid automated analysis of solution NMR spin-relaxation data. *J Biomol NMR* **26**, 203–213 (2003).
- Lipari, G. & Szabo, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *Journal of the American Chemical Society* **104**, 4546–4559 (1982).
- Lipari, G. & Szabo, A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *Journal of the American Chemical Society* **104**, 4559–4570 (1982).

27. Clore, G. M. *et al.* Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *Journal of the American Chemical Society* **112**, 4989–4991 (1990).
28. Tjandra, N., Kuboniwa, H., Ren, H. & Bax, A. Rotational dynamics of calcium-free calmodulin studied by 15N-NMR relaxation measurements. *Eur J Biochem* **230**, 1014–1024 (1995).
29. Loria, J. P., Rance, M. & Palmer, A. G. A Relaxation-Compensated Carr–Purcell–Meiboom–Gill Sequence for Characterizing Chemical Exchange by NMR Spectroscopy. *Journal of the American Chemical Society* **121**, 2331–2332 (1999).
30. Tollinger, M., Skrynnikov, N. R., Mulder, F. A., Forman-Kay, J. D. & Kay, L. E. Slow dynamics in folded and unfolded states of an SH3 domain. *J Am Chem Soc* **123**, 11341–11352 (2001).
31. Ishima, R. & Torchia, D. A. Estimating the time scale of chemical exchange of proteins from measurements of transverse relaxation rates in solution. *Journal of Biomolecular NMR* **14**, 369–372.
32. Cheung, M. S., Maguire, M. L., Stevens, T. J. & Broadhurst, R. W. DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J Magn Reson* **202**, 223–233 (2010).
33. Hwang, T. L., van Zijl, P. C. & Mori, S. Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical EXchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *J Biomol NMR* **11**, 221–226 (1998).
34. Rieping, W. *et al.* ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23**, 381–382 (2007).
35. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905–921 (1998).
36. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477–486 (1996).
37. Dempsey, C. E. Hydrogen exchange in peptides and proteins using NMR spectroscopy. *Progress in Nuclear Magnetic Resonance Spectroscopy* **39**, 135–170 (2001).
38. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **17**, 75–86 (1993).
39. Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43–56 (1995).
40. Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* (2013).
41. Berweger, C. D., van Gunsteren, W. F. & Müller-Plathe, F. Force field parametrization by weak coupling. Re-engineering SPC water. *Chemical Physics Letters* **232**, 429–436 (1995).
42. Lin, Z. & van Gunsteren, W. F. Refinement of the application of the GROMOS 54A7 force field to  $\beta$ -peptides. *Journal of Computational Chemistry* **34**, 2796–2805 (2013).
43. Schmid, N. *et al.* Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur Biophys J* **40**, 843–856 (2011).
44. Malde, A. K. *et al.* An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *Journal of Chemical Theory and Computation* **7**, 4026–4037 (2011).
45. Gaudreault, F., Morency, L. P. & Najmanovich, R. J. NRGsuite: a PyMOL plugin to perform docking simulations in real time using FlexAID. *Bioinformatics* **31**, 3856–3858 (2015).
46. Letourneau, D. *et al.* StAR-related lipid transfer domain protein 5 binds primary bile acids. *J Lipid Res* **53**, 2677–2689 (2012).
47. Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics* **81**, 511–519 (1984).
48. Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **31**, 1695–1697 (1985).
49. Henzler-Wildman, K. A. *et al.* A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913–916 (2007).
50. Kay, L. E. New Views of Functionally Dynamic Proteins by Solution NMR Spectroscopy. *J Mol Biol* **428**, 323–331 (2016).
51. Palmer, A. G. 3rd. Chemical exchange in biomacromolecules: past, present, and future. *J Magn Reson* **241**, 3–17 (2014).
52. Kudo, N. *et al.* Structural basis for specific lipid recognition by CERT responsible for nonvesicular trafficking of ceramide. *Proc Natl Acad Sci USA* **105**, 488–493 (2008).
53. Letourneau, D. *et al.* Thermodynamic and solution state NMR characterization of the binding of secondary and conjugated bile acids to STARD5. *Biochim Biophys Acta* (2013).
54. Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**, 1531–1534 (1996).
55. Yang, D. & Kay, L. E. Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding. *J Mol Biol* **263**, 369–382 (1996).
56. Forman-Kay, J. D. The ‘dynamics’ in the thermodynamics of binding. *Nat Struct Biol* **6**, 1086–1087 (1999).
57. Roy, J. & Laughton, C. A. Long-timescale molecular-dynamics simulations of the major urinary protein provide atomistic interpretations of the unusual thermodynamics of ligand binding. *Biophysical journal* **99**, 218–226 (2010).
58. Zidek, L., Novotny, M. V. & Stone, M. J. Increased protein backbone conformational entropy upon hydrophobic ligand binding. *Nat Struct Biol* **6**, 1118–1121 (1999).
59. Roderick, S. L. *et al.* Structure of human phosphatidylcholine transfer protein in complex with its ligand. *Nat Struct Biol* **9**, 507–511 (2002).
60. Baker, B. Y., Yaworsky, D. C. & Miller, W. L. A pH-dependent molten globule transition is required for activity of the steroidogenic acute regulatory protein, StAR. *J Biol Chem* **280**, 41753–41760 (2005).
61. Chang, I. Y. *et al.* A comparison of the steroidogenic acute regulatory protein-related lipid transfer (START) domain-containing 6 on the brain and testes between young and aged rats. *Acta Histochem* **116**, 551–558 (2014).
62. Gomes, C. *et al.* Expression of the putative sterol binding protein Stard6 gene is male germ cell specific. *Biol Reprod* **72**, 651–658 (2005).
63. Chang, I. Y. *et al.* Does the StarD6 mark the same as the StAR in the nervous system? *J Chem Neuroanat* **40**, 239–242 (2010).
64. Chang, I. Y., Kim, J. H., Cho, K. W. & Yoon, S. P. Acute responses of DNA repair proteins and StarD6 in rat hippocampus after domoic acid-induced excitotoxicity. *Acta Histochem* **115**, 234–239 (2013).
65. Chang, I. Y. *et al.* Immunohistochemical detection of StarD6 in the rat nervous system. *Neuroreport* **18**, 1615–1619 (2007).
66. Chang, I. Y. *et al.* The changed immunoreactivity of StarD6 after pilocarpine-induced epilepsy. *Neuroreport* **20**, 963–967 (2009).
67. Chang, I. Y. *et al.* Immunolocalization of steroidogenic acute regulatory protein-related lipid transfer (START) domain-containing proteins in the developing cerebellum of normal and hypothyroid rats. *J Chem Neuroanat* **43**, 28–33 (2012).
68. Tsutsui, K. Biosynthesis mode of action and functional significance of neurosteroids in the developing Purkinje cell. *J Steroid Biochem Mol Biol* **102**, 187–194 (2006).
69. Ukena, K., Kohchi, C. & Tsutsui, K. Expression and activity of 3 $\beta$ -hydroxysteroid dehydrogenase/delta5-delta4-isomerase in the rat Purkinje neuron during neonatal life. *Endocrinology* **140**, 805–813 (1999).
70. Tsutsui, K. Neurosteroids in the Purkinje cell: biosynthesis, mode of action and functional significance. *Mol Neurobiol* **37**, 116–125 (2008).

71. LaVoie, H. A. *et al.* STARD6 is expressed in steroidogenic cells of the ovary and can enhance *de novo* steroidogenesis. *Exp Biol Med (Maywood)* **239**, 430–435 (2014).

### Acknowledgements

This work was supported by a grant (MT-10983) to J.-G.L. and P.L. from the Canadian Institutes of Health Research. This work was also supported by the Regroupement Stratégique sur la Structure, la Fonction et l'Ingénierie des Protéines (PROTEO) and NSERC (P.L.).

### Author Contributions

D.L., J.-G.L. and P.L. conceived the study and wrote the paper. D.L. and A.L. produced and purified STARD6. D.L. ran and assigned the NMR spectra and the  $^{15}\text{N}$  relaxation experiments. D.L. also ran the amide exchange as well the binding and studies and analysed the data. D.L., M.B. and P.L. analyzed the NMR data. M.B. ran the structure calculations and the refinements. J.C. ran the MD simulations and analysed the trajectories. All the authors contributed to the edition of the manuscript and the writing of specific sections.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Létourneau, D. *et al.* STARD6 on steroids: solution structure, multiple timescale backbone dynamics and ligand binding mechanism. *Sci. Rep.* **6**, 28486; doi: 10.1038/srep28486 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>