The membrane anchor of the transcriptional activator SREBP is characterized by intrinsic conformational flexibility

Rasmus Linser[a,b,1], Nicola Salvi[b,c], Rodolfo Briones[d], Petra Rovó[a], Bert L. de Groot[d], and Gerhard Wagner[b,1]

[a]Department NMR-Based Structural Biology, Max-Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany; [b]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115; [c]Université Grenoble Alpes, Centre National de la Recherche Scientifique, and Commissariat à l'Énergie Atomique et aux Énergies Alternatives, Institut de Biologie Structurale, F-38044 Grenoble, France; and [d]Biomolecular Dynamics Group, Max-Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany

Regulated intramembrane proteolysis (RIP) is a conserved mechanism crucial for numerous cellular processes, including signaling, transcriptional regulation, axon guidance, cell adhesion, cellular stress responses, and transmembrane protein fragment degradation. Importantly, it is relevant in various diseases including Alzheimer's disease, cardiovascular diseases, and cancers. Even though a number of structures of different intramembrane proteases have been solved recently, fundamental questions concerning mechanistic underpinnings of RIP and therapeutic interventions remain. In particular, this includes substrate recognition, what properties render a given substrate amenable for RIP, and how the lipid environment affects the substrate cleavage. Members of the sterol regulatory element-binding protein (SREBP) family of transcription factors are critical regulators of genes involved in cholesterol/lipid homeostasis. After site-1 protease cleavage of the inactive SREBP transmembrane precursor protein, RIP of the anchor intermediate by site-2 protease generates the mature transcription factor. In this work, we have investigated the labile anchor intermediate of SREBP-1 using NMR spectroscopy. Surprisingly, NMR chemical shifts, site-resolved solvent exposure, and relaxation studies show that the cleavage site of the lipid-signaling protein intermediate bears rigid α-helical topology. An evolutionary conserved motif, by contrast, interrupts the secondary structure ~9–10 residues C-terminal of the scissile bond and acts as an inducer of conformational flexibility within the carboxyl-terminal transmembrane region. These results are consistent with molecular dynamics simulations. Topology, stability, and site-resolved dynamics data suggest that the cleavage of the α-helical substrate in the case of RIP may be associated with a hinge motion triggered by the molecular environment.

SREBP | regulated intramembrane proteolysis | membrane proteins | cellular lipid homeostasis | cancer

The expression of the cellular machinery responsible for cholesterol and fatty acid biosynthesis is tightly controlled by transcriptional regulation (1, 2). The negative-feedback loop, switching on transcription of target genes upon low fatty acid and cholesterol levels, is achieved via binding of transcriptional activators to the sterol regulatory element. This element is present in promoter regions near encoding genes (3, 4). Key players regulated include HMG-CoA reductase, the rate-limiting enzyme in cholesterol biosynthesis and target of statins, the low-density lipoprotein receptor, fatty acid synthase, and other enzymes vital for cellular lipid homeostasis. Consequently, aberrant regulation of these processes is linked to various conditions such as cardiometabolic disorders, inflammatory conditions, and aberrant development. Increased de novo fatty acid and cholesterol production has also been shown to be a hallmark of many human epithelial cancers (5, 6). Sterol regulatory element-binding proteins (SREBPs) are members of the bHLH-Zip transcription factor family using a basic helix–loop–helix–leucine zipper motif but are unique in that they are produced as inactive precursors anchored to the endoplasmic reticular (ER) membrane (3, 7). The SREBP precursor proteins consist of a two-helix membrane anchor, to which the amino (N)-terminal DNA-binding domain, including the transactivation domain (8), and a carboxyl (C)-terminal regulatory domain (CTD) are attached (Fig. 1A) (9). The extensive 31-residue hydrophilic loop linking the two transmembrane (TM) domains projects into the lumen of the ER. SREBP-2 has been shown to control the expression of genes governing the uptake and biosynthesis of cholesterol. Upon starvation of cholesterol, the SREBP-2 precursor is relocated from the ER to the Golgi by COPII vesicle transport. This involves dissociation of Insig1 from SCAP, which is bound to the CTD of SREBP. Site-1 protease (S1P) and S2P, located in the Golgi, in turn exert release of the active transcription factor (2). SREBP-1, which is linked to regulation of fatty acid, phospholipid, and triacylglycerol synthesis, can be activated by retrograde transport of the respective proteases from the Golgi into the ER (10). Cleavage of the SREBP precursor molecule arises in two steps. S1P separates the C-terminal TM helix, linked to the regulatory domain, from the N-terminal TM helix, tethered to the DNA-binding domain, at a cleavage site located within the long connecting loop (11). Subsequent cleavage by S2P occurs via regulated intramembrane proteolysis (RIP) (12) within the TM space to liberate the active transcription factor (see Fig. 1A).

Although the core domain of S2P has successfully been crystallized (13), details of the cleavage mechanism, such as domain mobility, substrate entry, and recognition, remain elusive to date. Despite a number of structures solved recently (14–17), these mechanisms

Significance

Sterol regulatory element-binding protein (SREBP) signaling is responsible for transcriptional regulation of cellular lipid homeostasis. Aberrant regulation of the feedback loop that switches on transcription of related genes upon depletion of fatty acids and cholesterol abrogates cellular integrity and is a hallmark of many cancers and other diseases. Regulated intramembrane proteolysis is a conserved mechanism also involved in other signaling processes like Wnt and Notch and defines cellular regulation from homeostasis to proliferation, compartmentalization, and differentiation. An understanding of basic features, however, such as substrate selectivity and specificity, has remained elusive to date. Understanding the proteolysis of the SREBP precursor is an important goal both for understanding and medically intervening in SREBP signaling and related diseases as well as for understanding other regulated intramembrane proteolysis-dependent processes.
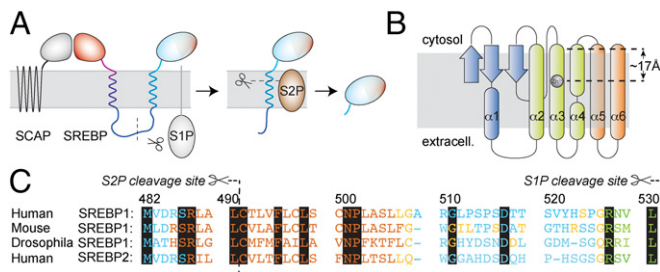
Fig. 1. The proteolytic cascade for maturation of SREBPs. (*A*) After COPII vesicle transport to the Golgi, the C-terminal part of the SREBP precursor, attached to SCAP, is shed by S1P. The active transcription factor (light blue) is then liberated by RIP of the single-TM S2P substrate. (*B*) S2P topology derived from Protein Data Bank ID code 3B4R (13). (*C*) Alignment of full-length SREBP anchor sequences. Key residues, which are conserved among different kingdoms and in between SREBP-1 and -2, are shown on black background. JPRED prediction (40) of protein TM regions (brown, helix and less than 25% accessible to solvent; blue, random coil and more than 25% solvent accessible; yellow, coil and less than 25% accessible; green, β-strand and less than 25% accessible). The region predicted as a β-strand obtains random-coil propensity when only the strip until the S1P cleavage site is subjected to the prediction (*SI Appendix*).

are poorly understood for the entire family of RIP proteases, including other important members such as presenilins (16, 18–20) and signal peptide peptidases (15, 21). Apart from regulating SREBP transcription factors, RIP has also been associated with several processes determining cell homeostasis and proliferation, involving Wnt and Notch signaling, β-catenin signaling, and cleavage of the amyloid precursor protein (APP) (22). Accordingly, dysfunction of these proteases plays a role in manifold diseases. All substrates of intramembrane proteases are assumed to be TM α-helices, even though protein helices are considered poor substrates for proteolysis (23). Protease substrates typically exhibit extended conformation (24, 25), and side chains bind to the active site for substrate recognition. Currently, it is largely unclear what makes a substrate suitable or unsuitable for RIP and how the lipid environment within the membrane space impinges on these properties (26, 27). For example, γ-secretase, involved in cleavage of the APP and the Notch receptor, has a large number of different substrates, for which only a loose definition of common features can be given (28, 29). In rhomboids, the membrane-integral serine protease relatives of S2Ps, which regulate epidermal growth factor receptor (EGFR) signaling in *Drosophila* (17, 30), common features for substrate recognition (31, 32) have long been elusive. They are now thought to involve a recognition motif and a helix-destabilizing motif (33), dependent on whether this and the cleavage site are inside the predicted TM helix (34, 35). A bound substrate-derived inhibitor cocrystalized with rhomboid (36) has indeed been found to be in an elongated β-strand form.

In the case of the S2P metalloproteases, a deeper understanding of the factors that make the SREBP membrane anchor a suitable target for recognition and enzymatic processing is still missing. Peptide bonds within regular α-helices are sterically hindered due to helix-stabilizing hydrogen bonds (37). Given its enthalpic stability, unfolding of a prototypical α-helix seems energetically unfavorable. Three SREBP sequence motives that are conserved and imperative for regulatory functionality have been described by Brown and Goldstein and coworkers (38): (*i*) The recognition site is constituted by a DRSR sequence just preceding the first predicted TM helix, (*ii*) the cleavage site is located at a Leu–Cys bond three residues into the hydrophobic segment (39), and (*iii*) an Asn–Pro motif is conserved in all SREBPs from worms to flies and humans at a position 11 amino acids distal to the cleavage site. Interestingly, replacement of any amino acid in the TM helix primary sequence with alanine does not alter cleavage of the helix in vivo (39). However, replacement of both N and P in the

conserved NP motif has been shown to abrogate S2P activity; on the other hand, this motif can be shifted up to five positions with little effect on cleavage activity. In work by Ye et al. (38), the region around this Pro has been suspected to induce a straightening of the helix and as such push the cleavage site out of the intramembrane space for proteolytic cleavage. After communication of the S2P X-ray structure (13), which bears an active center within the membrane space (Fig. 1*B*), this mechanism seems unlikely. Here we describe the topology and stability of the SREBP-1 membrane anchor to obtain information about RIP substrate characteristics as related to lipid regulation by S2P.

Results

We cloned the SREBP-1 membrane anchor in accordance with the conserved sequence and in line with the topology and membrane-integration predictions from JPRED (40) (depicted in Fig. 1*C*). Bacterial expression of SREBP-1 with or without fusion to different solubility-enhancing proteins in different *Escherichia coli* cell lines failed consistently, probably due to protease degradation of the partly unfolded peptides (41). Therefore, after extensive testing and optimization, SREBP-1 was eventually produced using cell-free expression and reconstituted in 1-myristoyl-2-hydroxy-*sn*-glycero-3-phospho-(1′-*rac*-glycerol) (LMPG) micelles (*SI Appendix, Details*). The SREBP-1 anchor topology is expected to differ from helix bundle membrane proteins, in which stable packing is required for physiological function. The SREBP-1 anchor contains no apparent packing motifs and is not predicted to pack by prediction routines. Instead the membrane anchor appears to consist of a combination of two individual membrane-embedded stretches, which are separated by a 31-amino-acid linker. This long luminal linker enables an independent function of the two TM helices after S1P action. Consequently, detergent screens pursued here (*SI Appendix*, Fig. S3) focused on the S2P substrate as truncated after S1P cleavage and were limited to folding, solubilization, and suitability for NMR characterization of this single-TM intermediate.

Solvent Accessibility in LMPG Micelles. Fig. 2*B* depicts the accessibility of the anchor residues to solvent as measured with amide-to-water nuclear Overhauser effect (NOE). The first assigned residue, R487, which is the last residue of the DRSR sequence, still shows clear solvent accessibility. The following stretch until six residues after the conserved NP motif shows no significant NOEs to water, whereas strong amide–amide NOEs are observed instead (see *NOEs*). Obvious water accessibility is again observed for the C-terminal linker residues. NOE cross-peaks to lipid hydrogens are found only at the interface between putative lipid-embedded and solvent-exposed residues. Interestingly, the C terminus also harbors some contacts to fatty acid hydrogens in addition to (more intense) water–NOE cross-peaks. This may be induced by contacts to the micelle or transient binding of single detergent molecules to these residues under in vitro conditions. The contacts point to some hydrophobicity of the five C-terminal residues and could impose an amphipathic effect of the truncated protein. The observation is in line with a temporary interaction of the C-terminal linker residues with the membrane in molecular dynamics (MD) simulations (see *Molecular Dynamics*). See *SI Appendix*, Fig. S7 for water/lipid NOE data on a construct including the C-terminal TM helix.

Secondary Chemical Shifts. Fig. 2*D* shows the secondary chemical shifts (SCSs) of C^α, C^β, and CO as obtained from sequential backbone assignment experiments as well as a TALOS-derived (42) secondary structure. For respective plots including the second helix, see *SI Appendix*, Figs. S7 and S8. From the first assigned residue, the second Arg of the DRSR motif, we observe a clear α-helical pattern. This continues through the S2P cleavage site until two residues before the conserved NP motif, where reduced helicity is found. The NP motif itself is unassignable due to the absence of the Pro H^N and the reduced viability of coherences from
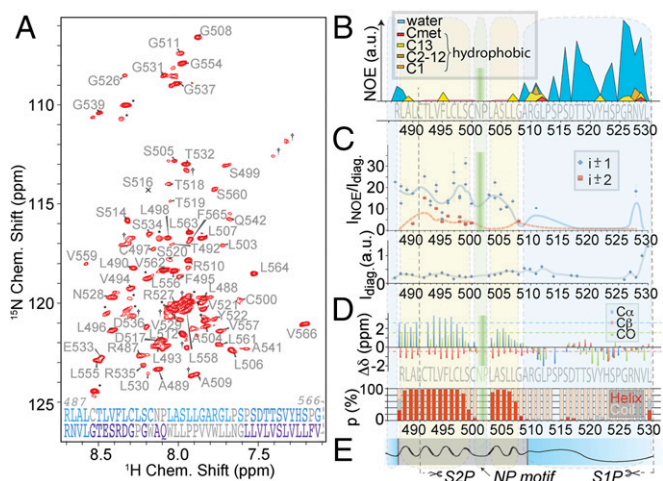
Fig. 2. Topology and putative membrane insertion of the SREBP-1 membrane anchor after S1P cleavage. (*A*) Two-dimensional proton–nitrogen correlation (TROSY) spectrum with assigned residues at 800 MHz ^1H Larmor frequency of the complete anchor (M482 to V566). Asterisks mark residues assigned but belonging to the uncleaved N-terminal tag. Crosses mark unassigned residues. The S516 peak is below the contours shown in the plot. (*Bottom*) Gray, light blue, and purple denote unassigned amino acids of the construct, those present, and those absent after S1P cleavage, respectively. (*B*) NOE contacts of amide protons to water (blue) and aliphatic (fatty acid) protons of the S2P substrate (M482 to L530). Hydrophobic contacts (yellow, orange, and red) are broken up into such methyl and methylene protons with approximate assignments within the chain. (*C*) Numbers, kind, and peak heights of NOE contacts observed for sequential ($i \pm 1$, blue) and $i \pm 2$ correlations (red) as a qualitative measure of perseverance of secondary structure. Blue (solid) and red (dashed) lines represent the trends (floating average) along the primary sequence for abundance of NOE transfer from direct and longer range contacts, respectively. NOE intensities were corrected for relaxation differences using diagonal-peak intensities shown below. The figure displays resolved peaks only. (*D*) Experimental SCS values represent differences to random coil values for C^α (blue), C^β (red), and CO shifts (green) averaged over a triple of adjacent amino acids. For the deuterated cell-free expression, deuterated amino acids Gln, Asn, Cys, and Trp were not available and were used in their protonated form, which often prohibits sufficient signal intensity for sequential assignments of these amino acids. The secondary structure in terms of TALOS+ prediction (42) is depicted below. (*E*) Topology model of the substrate (black line) concluded from the above data, representing helical (sinosoidal) and unstructured regions (stretched) as well as lipid (gray) and aqueous embedding (blue).

(undeuterated) Q, N, W, and C under slow-motional dynamics (*SI Appendix*). The three-residue stretch just after the NP motif shows increasing helicity propensities, which then decrease to very low values for another three amino acids. The remaining residues until the C terminus of the construct, as defined by the S1P cleavage site, possess a random-coil conformation.

NOEs. We characterized the amide-to-amide NOE patterns that the anchor residues show in ^{15}N-edited NOESY experiments. In line with secondary structural characterization by SCSs, we find that significant interresidue NOE intensities are present only in the stretch between R487 and G508. In addition, an $i - 1$ cross-peak is also observed for G511. We exclusively observe $i \pm 1$ and $i \pm 2$ contacts, as is expected for α-helical domains. Interestingly, also in combination with hydrophobicity and solvent accessibility of the C-terminal residues in the linker, N528 bears a sequential NOE cross-peak to R527 of significant intensity. This confirmed NOE is the only amide–amide contact observed for the solvent-exposed C-terminal part of the anchor. Fig. 2*C* shows the intensity and kind of NOE cross-peaks to each anchor residue as a function of residue. Intensity data were corrected by the height of the diagonal peak to account for differences in transfer efficiencies and line width. The intensity and number of contacts throughout the helical stretch

differ, however a qualitative trend can be derived by fitting the residue-specific NOE abundance over the sequence, using an interpolation of the running average. This approximate trend line, which is depicted in red/blue in the same plot as a qualitative measure of conformational stability, is in agreement with the SCSs described above. A lower degree of order is found at the C-terminal end of the lipid-embedded motif, commensurate to the picture obtained from SCSs. Lower structural stability is found directly next to the NP motif (with three comparably weak cross-peak totals in the two adjacent amino acids), which is in line with the helix-breaking tendency of Pro residues generally, and interestingly also around position F495. A lower helicity of this site is not suggested by the SCSs. However, low structural stability here is also apparent in relaxation experiments (see *Backbone Dynamics*). Diagonal-peak intensities used for computing the relative NOE cross-peak intensities are plotted in Fig. 2 *C*, *Bottom*. The high intensity at the C-terminal residues is due to little longitudinal and transverse relaxation (see also *Backbone Dynamics*) and is an indication of high mobility. The concluded anchor topology is represented in Fig. 2*E*.

SCSs as interpreted by TALOS (Fig. 2*D*) and NOE data were used to generate a topology model from CNS (43) (*SI Appendix*, Fig. S10). In contrast to the local structural features, however, no concise relative orientation of the upper and lower part of the helical elements could be obtained from paramagnetic relaxation enhancement (PRE) distance restraints (*SI Appendix*, Fig. S9) in micelles.

Backbone Dynamics. We pursued characterization of backbone dynamics using longitudinal (T_1) and transverse relaxation times (T_2) of $^{15}N^H$ as well as heteronuclear NOE (^1H, ^{15}N hetNOE). In addition, we acquired longitudinal and transverse CSA/dipole–dipole $^1H/^{15}N$ cross-correlated cross-relaxation rates η_z and η_{xy}. These experiments were recorded on the SREBP-1 membrane anchor as processed by S1P, reflecting the anchor dynamics relevant at the stage of S2P integral-membrane protease cleavage. hetNOE ratios (*SI Appendix*, Fig. S11*A*) follow a consistent global trend reflecting the decrease of effective correlation times toward the C-terminal end of the protein. Only a weak decrease in hetNOE values around the NP motif and toward the membrane/solvent interface is observed. The hydrophilic linker increases in flexibility at the C terminus reaching negative hetNOE effects, as is similar for small molecules and intrinsically disordered proteins (IDPs). Transverse and longitudinal $^{15}N^H$ relaxation (*SI Appendix*, Fig. S11 *B and C*) reflects the same overall trend, however with a higher sensitivity to local differences, particularly for T_2. Notably, steadily increasing T_2 times at the C terminus are opposed by initially decreasing T_1 times, which then revert to the same values found for the membrane-integral residues. This behavior resembles characteristics of unfolded polypeptides. Within the TM segment, transverse relaxation rates surpass low rates at F495, which is in line with low amide–amide NOE magnetization transfer (see *NOEs*). T_1 times for this residue are only marginally longer than for the remainder of the membrane-imbedded residues but still represent a local maximum also. We pursued reduced spectral density mapping to obtain the values of the spectral density function at zero, ^1H, and ^{15}N frequency [$J(0)$, $J(\omega_H)$, and $J(\omega_N)$] (Fig. 3 *A–C*). $J(0)$ thereby represents the effective correlation time at each protein site, if we assume the spectral density function to be a Lorentzian (44). Its average value amounts to ~8 ns, which is significantly lower than what would be expected for a globular membrane protein of this size in micelles (45). In addition to the highly flexible C terminus, the data confirm the presence of motion in the C-terminal part of the membrane-embedded anchor region. The $J(0)$ data as a function of residue point to increasing flexibility at the membrane:solvent interfaces and lower amplitude dynamics in the center. The region between the NP motif and the solvent-exposed region deviates from this bell and provides evidence for increased motility also within the
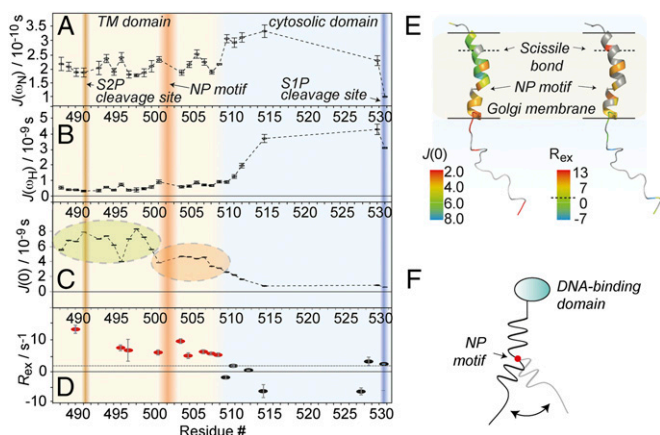
Fig. 3. Backbone dynamics of the membrane anchor as processed by S1P. (*A–C*) Spectral density mapping [$J(0)$, $J(\omega_H)$, and $J(\omega_N)$] derived from ^{1}H/^{15}N hetNOE, transverse and longitudinal ^{15}N autocorrelation rates R_2 and R_1, and transverse and longitudinal ^{1}H/^{15}N cross-relaxation rates η_z and η_{xy} data. Also see *SI Appendix*, Fig. S11. The S1P and S2P cleavage sites are marked by blue and brown shading, respectively; the NP motif is depicted in red. The background color of the plots denotes the position of the amino acids as membrane-imbedded (brown) or solvent-exposed (blue), according to Fig. 2. (*D*) Relaxation contributions from slow motion (µs to ns timescale motion). The dashed black line represents the average value. Residues with significant contributions from chemical exchange are depicted in red. (*E*) Model for the conformational flexibility derived from dynamics data, with the NP motif acting like a hinge. (*F*) Spectral density values at zero frequency (*Left*) and chemical-exchange contributions to R_2 (*Right*) depicted on a membrane anchor topology model (see *Molecular Dynamics* for details) as expected in a membrane environment. Residues with increased fast motional amplitudes would be marked by lower $J(0)$ (orange to red colors). The same colorization was used to mark significant exchange contributions (yellow to red colors). The membrane boundaries (as derived from water NOEs in micelles) are marked by the black lines; gray residues mark those with incomplete datasets for the determination for each parameter due to unclear assignments, overlapping H/N signals, or insufficient signal-to-noise ratio. Determination of relaxation parameters involves only dispersed peaks in all experiments. In addition, the Ser and Thr-rich solvent-exposed terminus is largely exchange-broadened below detection in the T_1 and T_2 relaxation experiments. The dashed line in the color legend represents the average R_{ex} throughout the protein.

membrane-embedded part. Here $J(0)$ values are halfway in between those in the more rigid N-terminal helical element and those in the neighboring flexible C-terminal loop region. Slightly enhanced motility is also found for the region around the conserved F495, which is six residues upstream of the NP motif. In addition to $J(0)$, also $J(\omega_H)$ and $J(\omega_N)$ values show that the C-terminal loop region is much more flexible than the helical element, but without providing any detail about the motional heterogeneity in the helix. The comprehensive relaxation dataset enabled us to determine the slow motion-derived (chemical exchange) contribution (µs to ms motions) to the ^{15}N transverse relaxation using the Kroenke approach (Fig. 3*D*) (46). These data provide evidence of slow-motional contributions to ^{15}N transverse relaxation within the putative membrane-embedded space, as opposed to the linker. Together with the $J(0)$ data, this may indicate that fast and slow motion coexist for part of the lipid-embedded stretch. Fig. 3*E* represents local relaxation properties mapped onto a topology model of the S1P-processed SREBP-1 membrane anchor as obtained after equilibrating the CNS model in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid environment by molecular dynamics (see *Molecular Dynamics*).

Molecular Dynamics. Using MD simulations, we equilibrated the SREBP-1 anchor structural model obtained from NMR data in micelles (*SI Appendix*, Fig. S10) into different lipid bilayers and monitored the secondary-structure stability over a period of up

to 1 µs as well as the backbone root-mean-square fluctuations (RMSFs) as a function of residue number (Fig. 4 *B* and *D* and *SI Appendix*, Fig. S12). In accordance with the NMR data, the stretch around the scissile bond remains α-helical and consistently rigid. The C-terminal helical element of the membrane-embedded part also retains its α-helical topology, and a hinge motion is observed with respect to the N-terminal helical element, again in full agreement with the NMR data. Remarkably, when we use membranes of lower thickness [1,2-dilauroyl-sn-glycero-3-phosphocholine (DLPC) instead of POPC lipids; Fig. 4*C*], we obtain an anchor with a steeper tilt angle (DLPC mean values of 48.3°, fluctuating between ~30 and 80°, opposed to a mean of 24.3°, with fluctuations only between 15 and 35°, for POPC) between the two helical elements, with the NP motif functioning as a hinge. A consistent trend from tilted conformations toward the more linear conformation (for detergent micelles via thin membranes to thick membranes) hints to a stretching effect of the anchor in response to the environment. In addition, slightly different amplitudes of the hinge motion (~17.7° and 8.7° for DLPC and POPC, respectively) are present in the two membranes (Fig. 4 *B* and *C* and *SI Appendix*, Fig. S13). Upon replacement of the lipids by water, fluctuations are significantly increased throughout the sequence, however complete unfolding is not observed within the timescale of the simulation (*SI Appendix*, Fig. S14). *SI Appendix*, Fig. S15 shows the structural model docked into the S2P structure (Protein Data Bank ID code 3B4R) (13). The open conformation is sufficiently wide to accommodate the substrate in its α-helical conformation. This protease–substrate combination is, however, artificial and has to be treated with extreme care.

Hydropathy. We asked if the distinct motional features of the anchor are associated with characteristics on the primary sequence level. We used the European Molecular Biology Open Software Suite (EMBOSS) Pepwindow framework to obtain information
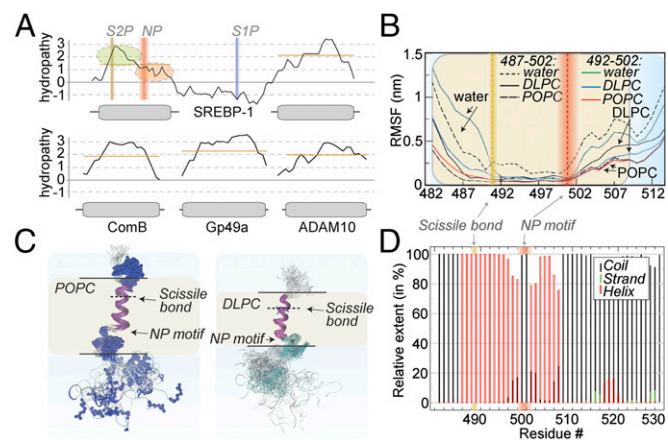

Fig. 4. Bioinformatics and molecular dynamics. (*A*) Hydropathy as a function of residue of SREBP-1 in comparison with other (non-protease substrate, single-TM helix) membrane proteins. The SREBP anchor is characterized by an asymmetric distribution of hydropathy values, implying very low hydropathy at the C-terminal end of the first TM helix. Orange bars depict the approximate average of hydropathy within each TM region. (*B*) RMSFs of the anchor in hydrophobic (DLPC and POPC membranes) and polar environment (exchange of lipids by water). Even in a polar environment, the C-terminal part of the helix shows a certain stability (also compare *SI Appendix*, Figs. S12 and S13). For alignment, residues 487–502 (black curves) or 492–502 (colored curves) were used. (*C*) Ensembles of equidistant snapshots in molecular dynamics trajectories over 1 µs. When accommodated in POPC (*Left*) and DLPC membranes (*Right*), the anchor displays significantly different tilt angles between the N-terminal and C-terminal part of the helix (see *Molecular Dynamics*). (*D*) Secondary-structural stability over a 1-µs molecular dynamics simulation in POPC. Random coil, β-strand, and α-helical structure is represented as black, green, and red bars, respectively.

about the hydropathy of the sequence. In fact, compared with a series of single-spanning membrane proteins (Fig. 4A), one obtains a strikingly asymmetric profile for the first helix of the SREBP-1 membrane anchor. In contrast to the N-terminal part of the helix, which accommodates the scissile bond, hydropathy values between the NP motif and the lipid:cytosol interface are very low and show a continuous, shallow slope. The shallow slope and low hydrophathy values in the C-terminal helical region of the S2P substrate are conserved among different kingdoms and between SREBP-1 and -2 (*SI Appendix*, Fig. S16).

Discussion

The characterization of the SREBP-1 membrane anchor both by NMR in micelles and by MD in lipid bilayers shows a topology consisting of two membrane-integral α-helical amino acid stretches separated by a flexible, solvent-accessible linker domain with no clear secondary structure. This linker was speculated in the work by Brown and Goldstein and coworkers (11) to be required for sufficient separability of the two helices for S1P action and to facilitate independent migration of the N-terminal part before S2P processing. In accordance with its considerably large expansion well beyond the common loop lengths in helix bundles, the C-terminal linker region indeed turns out to be highly mobile in our studies.

Peptide bonds within α-helices are poorly accessible to proteases because of the steric hindrance provided by the framework of helix-stabilizing hydrogen bonds (37). Accordingly, cytosolic proteases normally involve substrates in a β-strand extended conformation rather than relying on the poor cleavage properties of α-helices (24). Our data show that in the case of SREBP-1, similar to what is assumed for the other integral membrane proteases, the enzyme supposedly acts on a generally α-helical substrate. Interestingly, around the scissile peptide bond, the α-helical secondary structure even seems to be exceptionally stable, showing the lowest backbone flexibility throughout the primary sequence of the S2P substrate in terms of SCSs, NOEs, and relaxation data, and remains partly folded even upon switching the polarity of the environment in MD simulations. The length of the second helix (~29 residues) derived from the different experimental data are in the range of known single spanning TM helices. The first helix with only on the order of 22 nonsolvent accessible residues, however, involves significantly fewer residues within the membrane under the same conditions. This feature is in line with the imperfect α-helical topology derived from secondary chemical-shift data and an according stretched conformation around the NP motif. The stretching was originally hypothesized to push the scissile bond out of the hydrophobic space to allow protease cleavage (38), which seemed unlikely after finding the S2P active site within the hydrophobic space (13). Instead, the low helical SCSs, low numbers and intensities of sequential amide NOEs, and the local minimum of H/N hetNOE intensity at C500 just before the NP motif in this study strongly suggest this site to act as a helix-breaking motif that reduces local conformational stability. A high level of mobility, starting from just before the NP motif toward the C-terminal solvent interface, also within the membrane-spanning region, goes hand in hand with the interrupted helicity. This mobility, which is obvious from transverse and longitudinal relaxation rates in this element, with a local maximum at C500, is represented by a conformational flexibility as observed in MD simulations. Even though SCSs of this element clearly state a reversion to a loose α-helical topology for the remaining residues after the NP motif, the C-terminal integral-membrane part of the S2P substrate represents a structured but still dynamic element. Its dynamics increasing C-terminally and already approaching the flexibility of the solvent-exposed residues near the lipid–cytosol interface as seen by NMR are in line with a flexure of the element using the helix breaker as a hinge.

The lower stability of the C-terminal helical element is paired with a conserved amphipathic primary sequence, which contains the Arg of the NP motif and a Ser, two hydrophilic residues, and may

enable flexible accommodation in differentially hydrophobic environments. A tilt of this atypical element away from the remainder of the protein may be of relevance for a facilitated conformational adjustment for uptake by S2P and potentially differential unfolding upon contact with the protease. The bent conformation might also expose the polar Asn residue, which may be relevant for an initial interaction with S2P. Certainly, as derived from the different tilt angles in MD simulations of different membrane thickness, the flexure of the S2P substrate and all derived consequences will be directly influenced by the characteristics of the lipid membrane. Work by Walker et al. concluded from cell biological studies on *Caenorhabditis elegans* previously that membrane composition and curvature play a significant role for SREBP-1 cleavage in vivo, which could be related to the SREBP flexing (10).

In between the S2P cleavage site and the more mobile TM region toward the C terminus of the helix, F495 constitutes another site of enhanced mobility. For this region, we find significant contributions from slow-motional dynamics, as known from conformational preselection-type mechanisms. Whereas the functionally indispensable NP motif induces a required interruption of the α-helical conformation, the cleavage site seems completely buried before getting in contact with the protease and should require effective chaperone-like activity of the protease to unfold. This is in accordance with ample space in the S2P open conformation to accommodate the substrate in an α-helical (rather than extended) conformation. A similar picture has been obtained for APP, in which a flexible, bent TM domain might serve to most effectively interact with the protease (29). The TM region C-terminal to the hinge, however, is characterized by its more ambiguous hydropathy properties and the connection to the NP motif on one side and the highly mobile linker on the other. These destabilized residues in the C-terminal TM regions are thus likely to be unfoldable more easily in various conditions. Thus, apart from their tilt away from the NP motif, a partial unwinding of these residues might exist to further facilitate entry of the otherwise helical anchor into the protease. This could then exert any necessary unfolding of the internalized scissile bond by specific chaperone-like mechanisms.

For SREBP cleavage by S2P in the cell, quantitative details (dynamics time scales and amplitudes) and participants (constituting specific functions or crowding effects) will differ from the situation characterized under in vitro and in silico conditions. Still, the peculiar mechanical flexibility of the anchor is an intrinsic molecular property; thus, the observed flexure can well be expected to represent a significant feature of the substrate in vivo.

To summarize, the transcriptional activator SREBP regulates expression of machinery for cellular lipid and cholesterol homeostasis. A crucial step toward transcriptional activation, the liberation of the transcriptionally active regulatory domain, requires regulated intramembrane cleavage of the SREBP membrane anchor by S2P. Using solution NMR spectroscopy and molecular dynamics simulations, we find that the topology of this anchor is made up of a C-terminal regular TM helix, a long flexible linker, and an N-terminal TM helix. The substrate of the S2P cleavage, the anchor processed by S1P, bearing only the N-terminal TM domain and most of the linker, shows interrupted helicity at an evolutionarily conserved motif two-thirds down toward the linker. Whereas the clearly α-helical region bearing the scissile bond shows extraordinary rigidity, the lower elements of the TM domain, particularly the C-terminal helical element between the helix-breaking motif and the lipid-solvent interface, show reduced structural stability. The mobility of the residual linker outside the hydrophobic space increases to an extent comparable with IDPs C-terminally and is accompanied by fast-motional dynamics as well as chemical exchange within the destabilized TM region. Owing to the helix-breaking motif, which acts as a hinge in the S2P substrate, the substrate seems to be able to undergo conformational changes dependent on its molecular environment. In accordance with differential cleavage probabilities of SREBP-1 for different membranes reported previously, the conformational changes

observed in vitro and in silico may be relevant for the initiation and general feasibility of RIP cleavage.

Materials and Methods

The deuterated, $^{13}C/^{15}N$-labeled membrane anchor of the S2P substrate was expressed using cell-free expression. The purified protein was reconstituted into LMPG micelles. Samples were investigated by solution NMR at 800 MHz Larmor frequency. Established methodology was used for assignments and for characterization of topology and dynamics. MD simulations were performed in either POPC or DLPC lipid bilayers. Further details on the methods are provided in *SI Appendix*.

1. Brown MS, Goldstein JL (1998) Sterol regulatory element binding proteins (SREBPs): Controllers of lipid synthesis and cellular uptake. *Nutr Rev* 56(2 Pt 2):S1–S3; S54–S75.
2. Brown MS, Goldstein JL (2009) Cholesterol feedback: From Schoenheimer's bottle to Scap's MELADL. *J Lipid Res* 50(Suppl):S15–S27.
3. Brown MS, Goldstein JL (1997) The SREBP pathway: Regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* 89(3):331–340.
4. Eberlé D, Hegarty B, Bossard P, Ferré P, Foufelle F (2004) SREBP transcription factors: Master regulators of lipid homeostasis. *Biochimie* 86(11):839–848.
5. Menendez JA, Lupu R (2007) Fatty acid synthase and the lipogenic phenotype in cancer pathogenesis. *Nat Rev Cancer* 7(10):763–777.
6. Swinnen JV, Brusselmans K, Verhoeven G (2006) Increased lipogenesis in cancer cells: New players, novel targets. *Curr Opin Clin Nutr Metab Care* 9(4):358–365.
7. Brown MS, Goldstein JL (1998) Sterol regulatory element binding proteins (SREBPs): Controllers of lipid synthesis and cellular uptake. *Nutr Rev* 56(2 Pt 2):S1–S3; discussion S54–S75.
8. Párraga A, Bellsolell L, Ferré-D'Amaré AR, Burley SK (1998) Co-crystal structure of sterol regulatory element binding protein 1a at 2.3 A resolution. *Structure* 6(5): 661–672.
9. Nohturfft A, Brown MS, Goldstein JL (1998) Topology of SREBP cleavage-activating protein, a polytopic membrane protein with a sterol-sensing domain. *J Biol Chem* 273(27):17243–17250.
10. Walker AK, et al. (2011) A conserved SREBP-1/phosphatidylcholine feedback circuit regulates lipogenesis in metazoans. *Cell* 147(4):840–852.
11. Sakai J, et al. (1996) Sterol-regulated release of SREBP-2 from cell membranes requires two sequential cleavages, one within a transmembrane segment. *Cell* 85(7):1037–1046.
12. Brown MS, Ye J, Rawson RB, Goldstein JL (2000) Regulated intramembrane proteolysis: A control mechanism conserved from bacteria to humans. *Cell* 100(4):391–398.
13. Feng L, et al. (2007) Structure of a site-2 protease family intramembrane metalloprotease. *Science* 318(5856):1608–1612.
14. Hu J, Xue Y, Lee S, Ha Y (2011) The crystal structure of GXGD membrane protease FlaK. *Nature* 475(7357):528–531.
15. Li X, et al. (2013) Structure of a presenilin family intramembrane aspartate protease. *Nature* 493(7430):56–61.
16. Lu P, et al. (2014) Three-dimensional structure of human γ-secretase. *Nature* 512(7513): 166–170.
17. Urban S, Lee JR, Freeman M (2001) Drosophila rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell* 107(2):173–182.
18. Wolfe MS, et al. (1999) Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and γ-secretase activity. *Nature* 398(6727):513–517.
19. Struhl G, Greenwald I (1999) Presenilin is required for activity and nuclear access of Notch in Drosophila. *Nature* 398(6727):522–525.
20. De Strooper B, et al. (1999) A presenilin-1-dependent γ-secretase-like protease mediates release of Notch intracellular domain. *Nature* 398(6727):518–522.
21. Fluhrer R, Steiner H, Haass C (2009) Intramembrane proteolysis by signal peptide peptidases: A comparative discussion of GXGD-type aspartyl proteases. *J Biol Chem* 284(21):13975–13979.
22. Lal M, Caplan M (2011) Regulated intramembrane proteolysis: Signaling pathways and biological functions. *Physiology (Bethesda)* 26(1):34–44.
23. Loughlin WA, Tyndall JD, Glenn MP, Hill TA, Fairlie DP (2010) Update 1 of: Beta-strand mimetics. *Chem Rev* 110(6):PR32–PR69.
24. Tyndall JD, Fairlie DP (1999) Conformational homogeneity in molecular recognition by proteolytic enzymes. *J Mol Recognit* 12(6):363–370.
25. Madala PK, Tyndall JD, Nall T, Fairlie DP (2010) Update 1 of: Proteases universally recognize beta strands in their active sites. *Chem Rev* 110(6):PR1–PR31.
26. Lichtenthaler SF, Haass C, Steiner H (2011) Regulated intramembrane proteolysis—Lessons from amyloid precursor protein processing. *J Neurochem* 117(5):779–796.
27. Langosch D, Scharnagl C, Steiner H, Lemberg MK (2015) Understanding intramembrane proteolysis: From protein dynamics to reaction kinetics. *Trends Biochem Sci* 40(6):318–327.
28. Haapasalo A, Kovacs DM (2011) The many substrates of presenilin/γ-secretase. *J Alzheimers Dis* 25(1):3–28.
29. Barrett PJ, et al. (2012) The amyloid precursor protein has a flexible transmembrane domain and binds cholesterol. *Science* 336(6085):1168–1171.
30. Lee JR, Urban S, Garvey CF, Freeman M (2001) Regulated intracellular ligand transport and proteolysis control EGF signal activation in Drosophila. *Cell* 107(2):161–171.
31. Urban S, Schlieper D, Freeman M (2002) Conservation of intramembrane proteolytic activity and substrate specificity in prokaryotic and eukaryotic rhomboids. *Curr Biol* 12(17):1507–1512.
32. Koide K, Ito K, Akiyama Y (2008) Substrate recognition and binding by RseP, an Escherichia coli intramembrane protease. *J Biol Chem* 283(15):9562–9570.
33. Urban S, Freeman M (2003) Substrate specificity of rhomboid intramembrane proteases is governed by helix-breaking residues in the substrate transmembrane domain. *Mol Cell* 11(6):1425–1434.
34. Strisovsky K, Sharpe HJ, Freeman M (2009) Sequence-specific intramembrane proteolysis: Identification of a recognition motif in rhomboid substrates. *Mol Cell* 36(6): 1048–1059.
35. Strisovsky K (2013) Structural and mechanistic principles of intramembrane proteolysis—Lessons from rhomboids. *FEBS J* 280(7):1579–1603.
36. Zoll S, et al. (2014) Substrate binding and specificity of rhomboid intramembrane protease revealed by substrate-peptide complex structures. *EMBO J* 33(20):2408–2421.
37. Paetzel M, Dalbey RE, Strynadka NCJ (1998) Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. *Nature* 396(6707):186–190.
38. Ye J, Davé UP, Grishin NV, Goldstein JL, Brown MS (2000) Asparagine-proline sequence within membrane-spanning segment of SREBP triggers intramembrane cleavage by site-2 protease. *Proc Natl Acad Sci USA* 97(10):5123–5128.
39. Duncan EA, Davé UP, Sakai J, Goldstein JL, Brown MS (1998) Second-site cleavage in sterol regulatory element-binding protein occurs at transmembrane junction as determined by cysteine panning. *J Biol Chem* 273(28):17801–17809.
40. Cole C, Barber JD, Barton GJ (2008) The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* 36(Web Server issue):W197–W201.
41. Linser R, et al. (2014) Selective methyl labeling of eukaryotic membrane proteins using cell-free expression. *J Am Chem Soc* 136(32):11308–11310.
42. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44(4):213–223.
43. Brünger AT, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54(Pt 5): 905–921.
44. Salvi N, Ulzega S, Ferrage F, Bodenhausen G (2012) Time scales of slow motions in ubiquitin explored by heteronuclear double resonance. *J Am Chem Soc* 134(5): 2481–2484.
45. Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry* 28(23):8972–8979.
46. Kroenke CD, Loria JP, Lee LK, Rance M, Palmer AG (1998) Longitudinal and transverse 1H-15N dipolar 15N chemical shift anisotropy relaxation interference: Unambiguous determination of rotational diffusion tensor and chemical exchange effect in biological macromolecules. *J Am Chem Soc* 120(31):7905–7915.

BIOPHYSICS AND COMPUTATIONAL BIOLOGY