

Communication

ReSMAP: Web Server for Predicting Residue-Specific Membrane-Association Propensities of Intrinsically Disordered Proteins

Sanbo Qin¹, Alan Hicks¹, Souvik Dey¹ , Ramesh Prasad¹ and Huan-Xiang Zhou^{1,2,*} ¹ Department of Chemistry, University of Illinois at Chicago, Chicago, IL 60607, USA² Department of Physics, University of Illinois at Chicago, Chicago, IL 60607, USA

* Correspondence: hzhou43@uic.edu

Abstract: The functional processes of many proteins involve the association of their intrinsically disordered regions (IDRs) with acidic membranes. We have identified the membrane-association characteristics of IDRs using extensive molecular dynamics (MD) simulations and validated them with NMR spectroscopy. These studies have led to not only deep insight into functional mechanisms of IDRs but also to intimate knowledge regarding the sequence determinants of membrane-association propensities. Here we turned this knowledge into a web server called ReSMAP, for predicting the residue-specific membrane-association propensities from IDR sequences. The membrane-association propensities are calculated from a sequence-based partition function, trained on the MD simulation results of seven IDRs. Robustness of the prediction is demonstrated by leaving one IDR out of the training set. We anticipate there will be many applications for the ReSMAP web server, including rapid screening of IDR sequences for membrane association.

Keywords: membrane binding; intrinsically disordered proteins; membrane-association propensity; amphipathic helix; intrinsically disordered regions



Citation: Qin, S.; Hicks, A.; Dey, S.; Prasad, R.; Zhou, H.-X. ReSMAP: Web Server for Predicting Residue-Specific Membrane-Association Propensities of Intrinsically Disordered Proteins. *Membranes* **2022**, *12*, 773. <https://doi.org/10.3390/membranes12080773>

Academic Editor: Hana Popelka

Received: 14 July 2022

Accepted: 8 August 2022

Published: 11 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The functional mechanisms of many proteins involve the association of their intrinsically disordered regions (IDRs) with acidic membranes [1]. For example, the Wiskott–Aldrich Syndrome protein (WASP) and its neuronal homologue (N-WASP) are autoinhibited until activated in part by the binding of a disordered basic region with the acidic lipid phosphatidylinositol (4,5)-bisphosphate (PIP₂) in the plasma membrane, leading to the release of C-terminal domains and the initiation of actin polymerization [2]. By attaching its basic domain to the plasma membrane and sequestering PIP₂, the disordered protein myristoylated alanine-rich C-kinase substrate maintains a PIP₂ reservoir, which can be released by binding with calmodulin [3]. The disordered C-terminal domains of the tetrameric ligand-gated ion channel protein NMDA receptor modulate its gating properties [4], likely through membrane association of the membrane-proximal regions. A number of membrane proteins that make up the cell division machinery of *Mycobacterium tuberculosis* have disordered cytoplasmic regions, whose membrane association mediates protein–protein interactions [5], particularly through the reduction in dimensionality. Protein–protein colocalization and interactions mediated by membrane association have also been implicated for the disordered intracellular domains of the prolactin receptor, the growth hormone receptor [6], and the T-cell receptor [7], as well as for the membrane-proximal domain of the sheddase ADAM17 [8]. SepF, a water-soluble protein in *M. tuberculosis* and other bacteria, tethers its N-terminal amphipathic helix to the inner membrane, allowing it to act as a membrane anchor for the Z-ring at the start of the cell division process [9]. Membrane targeting of the Src family of kinases is achieved in part by basic residues in the

disordered *N*-terminal region [10,11]. Both synaptobrevin-2 and α -synuclein have been suggested to promote membrane fusion by associating with membranes [12,13]. Knowledge of the residue-specific membrane association of these and many other IDRs (or intrinsically disordered proteins) will provide deep mechanistic insight into their functional processes.

Characterizing the residue-specific membrane association of IDRs presents significant challenges. NMR spectroscopy can provide the most detailed information, as done for a few IDRs [5,6,11–13]. Molecular dynamics (MD) simulations have recently become accurate for modeling the membrane association of IDRs but, even with GPU acceleration, months of simulation time may be required to cover their vast conformational space [5]. A fast method, such as one based on IDR sequences, is highly desirable. For association with acidic membranes, the importance of basic residues has been a universal observation, whereas the roles of aromatic and hydrophobic residues seem to be context-dependent [1,3].

Recently a sequence-based method has been developed for a related problem, i.e., for predicting the propensities of IDRs binding to nanoparticles [14]. The nanoparticle-binding propensities are predicted from a partition function that is determined by the IDR sequence. Here we adapt this method into a predictor called ReSMAP, for residue-specific membrane-association propensities.

2. Computational Methods

The training data for ReSMAP were obtained from molecular dynamics (MD) simulations of seven intrinsically disordered regions (IDRs): *N*-terminal regions of ChiZ, FtsQ, and SepF, which are components of the cell division machinery of *M. tuberculosis*; the membrane-proximal regions in the C-terminal domains of the GluN1 and GluN2B subunits of the NMDA receptor; and disordered regions of N-WASP and WASP. In addition, a separate IDR in SepF was simulated and the data were used solely for testing ReSMAP. The two SepF IDRs, spanning residues 1–50 and 66–124, are referred to as SepF1 and SepF2, respectively. The numbers of residues in these IDRs are 64, 99, 50, 105, 123, 85, 82, and 59, respectively (see Table S1 for sequences). The MD simulation protocol was as described by Hicks et al. [5]. The force field combination was AMBER14SB [15] for proteins, TIP4P-D for water [16], and Lipid17 [17] for membranes. The lipid compositions in the simulations of the eight IDRs are listed in Table S2. The total MD simulation times were 38 μ s for ChiZ (among 20 replicate runs), 1 μ s each for GluN1, GluN2B, and SepF1 (among 4 replicate runs), 16 μ s each for N-WASP, WASP, and FtsQ (among 16 replicate runs), and 2.88 μ s for SepF2 (among 8 replicate runs). Except for ChiZ [5], the simulation results were not reported previously and their functional implications will be reported elsewhere.

Previously we measured the level of membrane association by using the contact probability of each residue, i.e., the fraction of MD snapshots in which this residue forms at least one contact, defined with a 3.5 Å cutoff, between its heavy atoms and lipid heavy atoms [5]. Here we introduce another measure, based on z_{tip} , the mean z coordinate of the tip atom of each side chain in a Cartesian coordinate system where the xy plane is located on the phosphate plane of the membrane. We convert z_{tip} into a contact probability via a smooth function,

$$C = \frac{1}{1 + \exp[(z_{\text{tip}} - z_0)/L(z_{\text{tip}})]} \quad (1a)$$

where

$$L(z_{\text{tip}}) = L_1 - \frac{L_1 - L_0}{1 + \exp[(z_{\text{tip}} - L_m)/L_w]} \quad (1b)$$

Fit to the previous “raw” contact probability yields $z_0 = 6.9$ Å, $L_1 = 3.2$ Å, $L_0 = 1.2$ Å, $L_m = 5.0$ Å, and $L_w = 0.5$ Å (Figure S1).

Given the universal importance of basic residues in the association with acidic membranes, we first checked whether a toy model, based on the moving average charge calculated over a seven-residue window, had any merit. The charge of each amino acid was +1 (K, R, and the *N*-terminus), −1 (D, E, and the *C*-terminus), or 0 (all other). Even this toy model showed some promise (Figure S2): a linear regression analysis against the MD

membrane-contact probabilities yielded coefficients of determination (R^2) in the range of 0.20 to 0.58 for six fully disordered IDRs. Encouraged by the promising sign of the toy model, we sought a more sophisticated method for predicting residue-specific membrane-association propensities by following the basic idea behind the sequence-based partition function of Li et al. [14]. The modifications from the toy model are threefold. First, instead of an abrupt cutoff beyond a sequence distance of 3 residues when considering the effects of neighboring residues, we model the effects with a continuous function that attenuates with increasing sequence distance. Second, instead of an additive term, the effect of each neighboring residue is modeled as a multiplicative factor. Third, the model in theory allows for the amino acids to be divided into more than just three groups (positively charged; negatively charged; and neutral).

Specifically, we assume that the central residue, with index n , and all other residues in the sequence each contribute a multiplicative Boltzmann factor to the statistical weight for residue n 's membrane association:

$$w_n = \prod_{i=1}^N q_{i;|i-n|} \quad (2)$$

where the residue index i runs through the entire sequence (a total of N residues), and $q_{i;|i-n|}$ is the contributing factor of residue i . The latter depends on the amino-acid type of residue i and the sequence distance $|i - n|$:

$$q_{i;|i-n|} = 1 + \frac{q_{i;0} - 1}{1 + a_i|i - n| + b_i|i - n|^2} \quad (3)$$

The dependence on sequence distance is to attenuate the contributing factor as residue i moves farther from residue n along the sequence. Note that a residue with an amplitude $q_{i;0} > 1$ increases w_n whereas a residue with $q_{i;0} < 1$ decreases w_n . At present we only distinguish three types of amino acids: positively charged (K, R, and the N-terminus), negatively charged (D, E, and the C-terminus), and neutral (all other amino acids). We denote the $q_{i;0}$ values of these three types as q_+ , q_- , and q_0 , respectively. The a_i and b_i values are the same for all charged residues and are denoted by a_{\pm} and b_{\pm} , respectively. For uncharged amino acids, $b_i = 0$ and the common a_i value is denoted by a_0 . We inherit distance parameter values from Li et al. [14], with $a_{\pm} = 0.0982$, $b_{\pm} = 0.00305$, and $a_0 = 0.521$. The amplitude parameters q_+ , q_- , and q_0 are optimized against MD data for contact probabilities. The membrane-association propensity is proportional to w_n :

$$P_n = c \frac{w_n}{w_{\max}} \quad (4)$$

where w_{\max} is the maximum of w_n in the entire sequence, and c is a scaling factor (a constant for each protein).

3. Results

As reported previously, 13 positively charged R residues in the ChiZ IDR drive its association with acidic membranes [5]. The association is highly dynamic: each moment a different subset of the R residues forms membrane contact (see Figure 1a for a snapshot). The IDRs of ChiZ, GluN1, GluN2B, N-WASP, and WASP remain fully disordered while associating with acidic membranes. Upon optimizing the three parameters q_+ , q_- , and q_0 , Equation (4) accurately predicts the residue-specific contact probabilities for all these IDRs (Figures 1b,c and S3a–c). The parameter values are $q_+ = 2.43$, $q_- = 0.26$, and $q_0 = 0.59$ (Table S3). Only q_+ is > 1 , and so only positively charged residues favor membrane association; both negatively charged and uncharged residues disfavor membrane association, strongly in the former case and weakly in the latter case. The root-mean-square-errors (RMSEs) measured against the MD membrane-contact probabilities are 0.043, 0.035, 0.012, 0.023, and 0.011 for the IDRs in ChiZ, GluN1, GluN2B, N-WASP, and WASP, respectively.

These RMSEs are up to 3-fold lower than the counterparts produced by the linear regression equation using the moving average charge (Figure S2). As another measure of accuracy, we display the Equation (4) predictions and the corresponding MD contact probabilities as a scatter plot in Figure S4a. The points are all close to the diagonal line $y = x$, with a coefficient of determination (R_0^2) at a high value of 0.91. Note that R_0^2 essentially measures the RMSEs against the mean amplitudes of the MD contact probabilities.

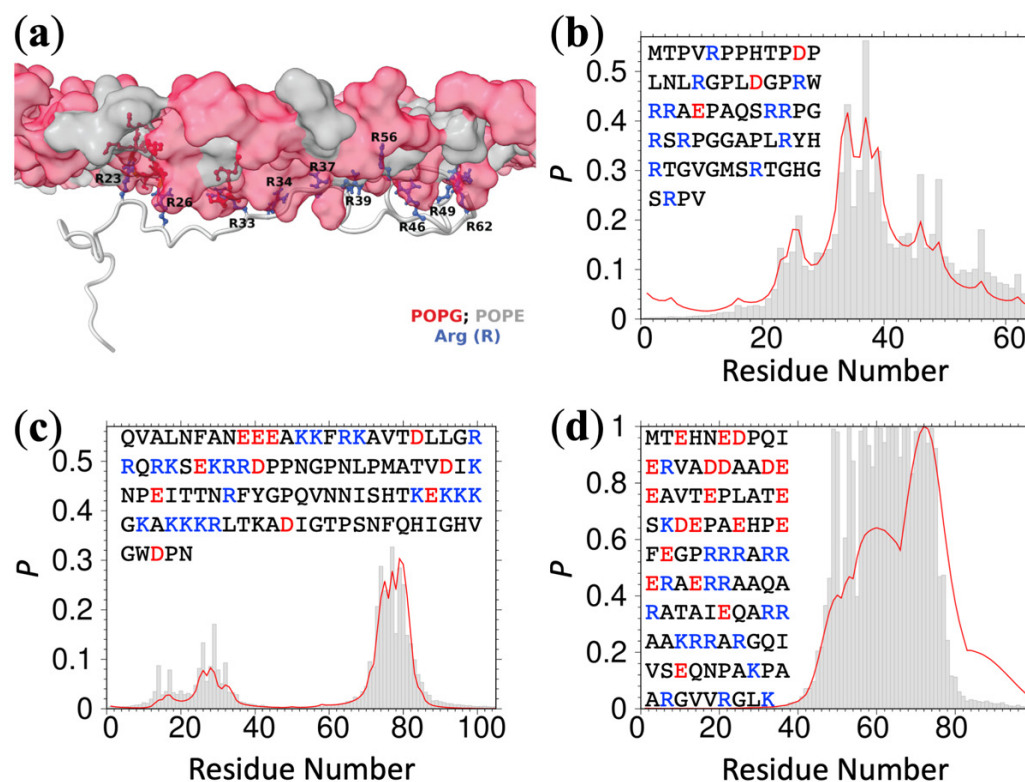


Figure 1. MD data and ReSMAP predictions for membrane-association propensities. (a) A snapshot of the ChiZ IDR associated to an acidic membrane, reprinted from Ref. [5]. The lipid headgroups in the leaflet in contact with the IDR are shown as surface; Arg side chains are shown as stick. (b–d) Comparison of MD membrane-contact probabilities (gray bars) and predicted membrane-association propensities (red curves) for ChiZ, N-WASP, and FtsQ, respectively. The sequence of each IDR is listed, with positively and negatively charged residues colored blue and red, respectively.

To demonstrate the robustness of the predictions, we compare the parameter values when one of the IDRs is left out of the training set. Because two of the IDRs, from N-WASP and WASP, have moderate sequence identity (34%), we left them out together in this exercise. The resulting values are 2.7 ± 0.7 (mean \pm standard deviation) for q_+ , 0.28 ± 0.08 for q_- , and 0.59 ± 0.04 for q_0 , agreeing well with those from the full training set. As one more test, we carried out MD simulations for another fully disordered IDR, SepF2. This IDR is not in the training set but the prediction, with an RMSE of 0.18, also agrees well with the MD results (Figure S3d). The R_0^2 value for this test IDR is still high, at 0.73 (Figure S4b).

Two other IDRs, FtsQ and SepF1, each have an amphipathic helix (residues 48 to 73 in FtsQ and 1–11 in SepF1) that stably associates with acidic membranes. The first 100 residues of α -synuclein also form amphipathic helices that stably associate with PIP₂, as characterized recently by NMR spectroscopy [13]. We combined the MD contact data for the FtsQ and SepF1 IDRs and the NMR data for every third residue of α -synuclein to optimize Equation (4). With $q_+ = 2.29$, $q_- = 0.64$, and $q_0 = 1.17$ (Table S3), the membrane-contact probabilities of the FtsQ and SepF1 IDRs and the entire α -synuclein sequence are predicted well (Figures 1d and S5a,b). The values of q_+ , q_- , and q_0 are essentially unchanged when other one thirds of the α -synuclein NMR data are used for training.

The RMSEs are 0.19, 0.27, and 0.17; the R_0^2 value, which as noted above measures the RMSEs against the mean amplitudes of the observed contact probabilities, is a high 0.83 (Figure S4c). The prediction accuracy is lower than that for fully disordered IDRs, possibly indicating that the sequence-based partition function, with only three adjustable parameters, is too restrictive for IDRs that have both amphipathic helices and disordered residues (see Discussion). Note that q_0 is now above 1 (albeit only slightly), likely due to the occurrence of neutral residues in amphipathic helices.

We further tested the foregoing ReSMAP method against residue-specific membrane association data of IDRs that we could find in published NMR studies. Chemical shift differences ($\Delta\delta_{\text{NH}}$) obtained by Haxholm et al. [6] suggested three lipid interaction sites at the *N*-terminus, middle, and *C*-terminus of the intracellular domain (residues G236-H598) of the prolactin receptor. ReSMAP predicts the *N*- and *C*-terminal sites and a middle site that partially overlaps with the NMR data (Figure 2a). In addition, it predicts high membrane-association propensities for the stretch of residues 509–512, with sequence KPKK. The same NMR study also revealed a single lipid interaction site, at the *N*-terminus of the intracellular domain (S270-P620) of the growth hormone receptor. ReSMAP predicts the same site (Figure 2b). Pond et al. [11] reported chemical shift differences for the disordered *N*-terminal region (G2-E79) of the Src kinase Hck, showing a membrane interaction site in residues R18-T40. Our ReSMAP prediction agrees well with this interaction site (Figure 2c). It further predicts high membrane-association propensities for the first three residues (GGR), which might be consistent with the NMR data since resonances of these residues appear to be missing. Missing resonances can result from membrane association. Lakomek et al. [12] studied the membrane association of the soluble portion (residues M1-M96) of the SNARE protein synaptobrevin-2 by NMR. The *C*-terminal 15 or so residues have a high propensity for forming a helix in solution. These residues also show strong binding with liposomes containing 20% DOPS, presumably as an amphipathic helix, as evidenced by missing NMR peaks; the membrane-bound population (as monitored by the ratio of peak intensities in the presence of liposomes and in solution) rapidly decreases toward the *N*-terminus. The ReSMAP prediction shows an identical trend (Figure 2d).

We also tested ReSMAP against experimental data for two other IDRs. Using in-cell FRET, Zhang et al. [7] monitored the binding of the intracellular domain of the T-cell receptor ζ chain to the plasma membrane. Alanine mutations of basic residues in three basic-rich stretches abolished membrane binding, whereas phenylalanine mutations of tyrosines in the so-called intracellular immunoreceptor tyrosine-based activation motifs had no effect. Consistent with these mutation results, ReSMAP predicts high membrane-association propensities for all the three basic-rich stretches and low membrane-association propensities for all the three intracellular immunoreceptor tyrosine-based activation motifs (Figure 2e). Lastly, Sommer et al. [8] identified a basic motif (R625-K626-G627-K628) for mediating interactions of the membrane-proximal domain (F581-E642) of ADAM17 with phosphatidylserine lipids, as indicated by significant chemical shift perturbations in the presence of phosphorylserine. Glycine mutations of the three basic residues abolished binding with phosphatidylserine liposomes. ReSMAP predicts a single high peak for the experimentally identified basic motif (Figure 2f). Together, the experimental data from the six proteins provide mounting evidence in support of the ReSMAP method.

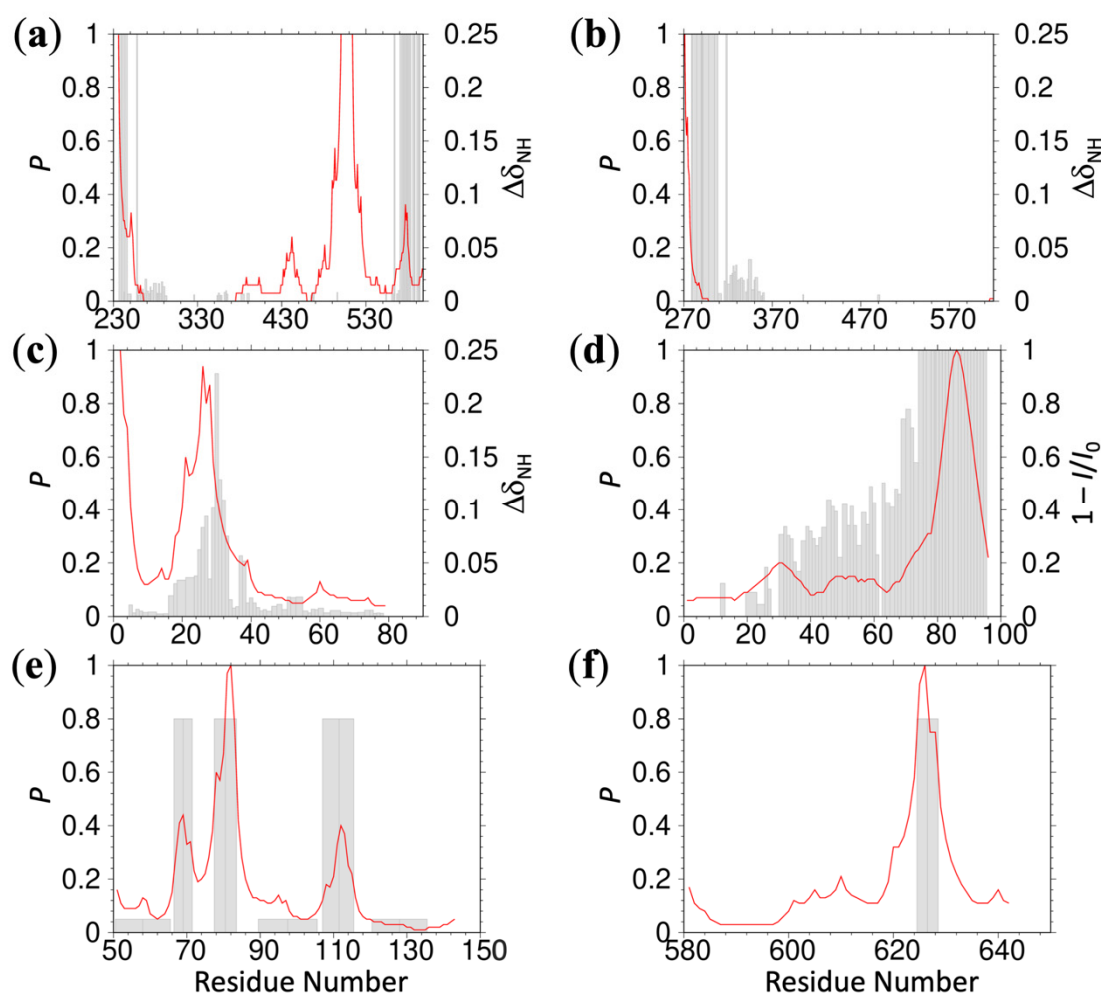


Figure 2. Comparison of ReSMAP-predicted membrane-association propensities (red curves) with experimental data (gray bars). (a,b) Prolactin receptor and growth-hormone receptor disordered intracellular domains. The experimental data are from Haxholm et al. [6], displaying $\Delta\delta_{\text{NH}} = \sqrt{(\Delta\delta_{\text{H}})^2 + 0.154(\Delta\delta_{\text{N}})^2}$, where $\Delta\delta_{\text{N}}$ and $\Delta\delta_{\text{H}}$ are changes in backbone amide chemical shifts when the IDRs were moved from solution to vesicles formed by POPC/POPS lipids at a 3:1 ratio. $\Delta\delta_{\text{NH}}$ values < 0.005 ppm were assumed to be within experimental error and set to 0; missing resonances were assigned a $\Delta\delta_{\text{NH}}$ value of 0.25 ppm. The protein sequences are from Uniprot (<https://www.uniprot.org/>, accessed on 28 July 2022) entries P16471 and P10912, with numbering shortened by 24 and 18 residues (i.e., without the *n*-terminal signal peptides), respectively. (c) Hck kinase disordered *N*-terminal region (residues G2-E79). The experimental data are from Pond et al. [11], displaying $\Delta\delta_{\text{NH}} = \sqrt{(\Delta\delta_{\text{H}})^2 + 0.2(\Delta\delta_{\text{N}})^2}$ obtained when the IDR was moved from solution to bicelles formed with DMPC:DMPA lipids at a 4:1 ratio. (d) Synaptobrevin-2 residues M1-M96. For the experimental data [12], I_0 and I were measured, respectively, in solution and in the presence of liposomes formed by DOPC/DOPS/DOPE/Cholesterol at 5:2:2:1. Resonances missing in the presence of liposomes were assigned a zero value for I . (e) T-cell receptor ζ chain disordered intracellular domain. The protein sequence is from Uniprot entry P24161, with numbering shortened by 21 residues. The tall bars indicate basic-rich stretches where alanine mutations of basic residues abolished plasma membrane binding, as assayed by in-cell FRET [7]. The short bars indicate tyrosine-containing motifs where phenylalanine mutations of tyrosine residues had no effect on plasma membrane binding. (f) ADAM17 membrane-proximal domain (residues F581-E642). Bars indicate four residues (R625-K628) that experienced significant chemical shift perturbations in the presence of phosphorylserine [8].

We have implemented ReSMAP as a webserver, accessible at <https://pipe.rcc.fsu.edu/ReSMAPidp/>. As an illustration of its application, a number of other components of the cell division machinery of *M. tuberculosis*, including FtsB and FtsL, are also likely to associate with acidic membranes via IDRs. We present the predicted membrane-association propensities of these IDRs in Figures 3 and 4.

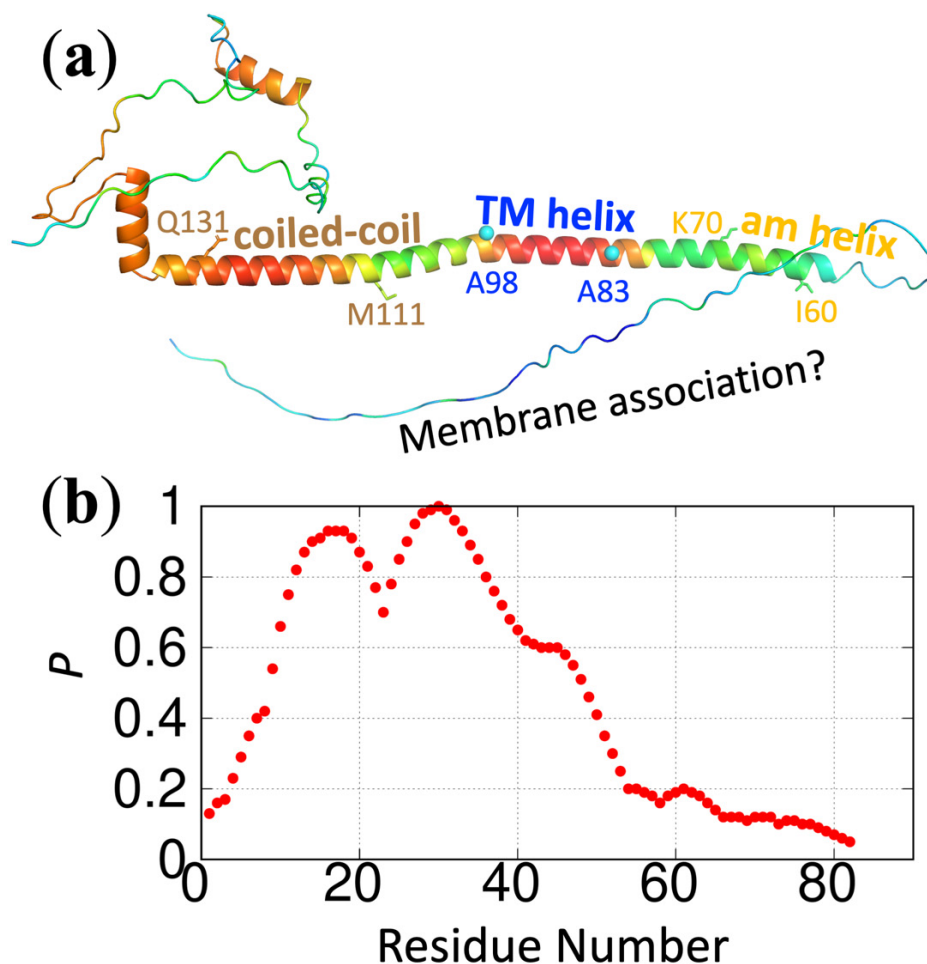


Figure 3. Predicted membrane-association propensities of FtsB. (a) Predicted structure by AlphaFold (<https://alphafold.ebi.ac.uk/>, accessed on 18 March 2022) [18]. The color spectrum displays confidence levels of prediction, ranging from red for high confidence to blue for low confidence. Three putative helices are indicated with start and end residues: amphipathic (am) helix, residues I60-K70; transmembrane (TM) helix, residues A83-A98; and coiled-coil, M111-Q131. (b) Predicted membrane-association propensities using the sequence of the first 82 residues, high for residues 9–48 but only modest for the putative amphipathic helix. The protein sequence is from Uniprot entry P96376.

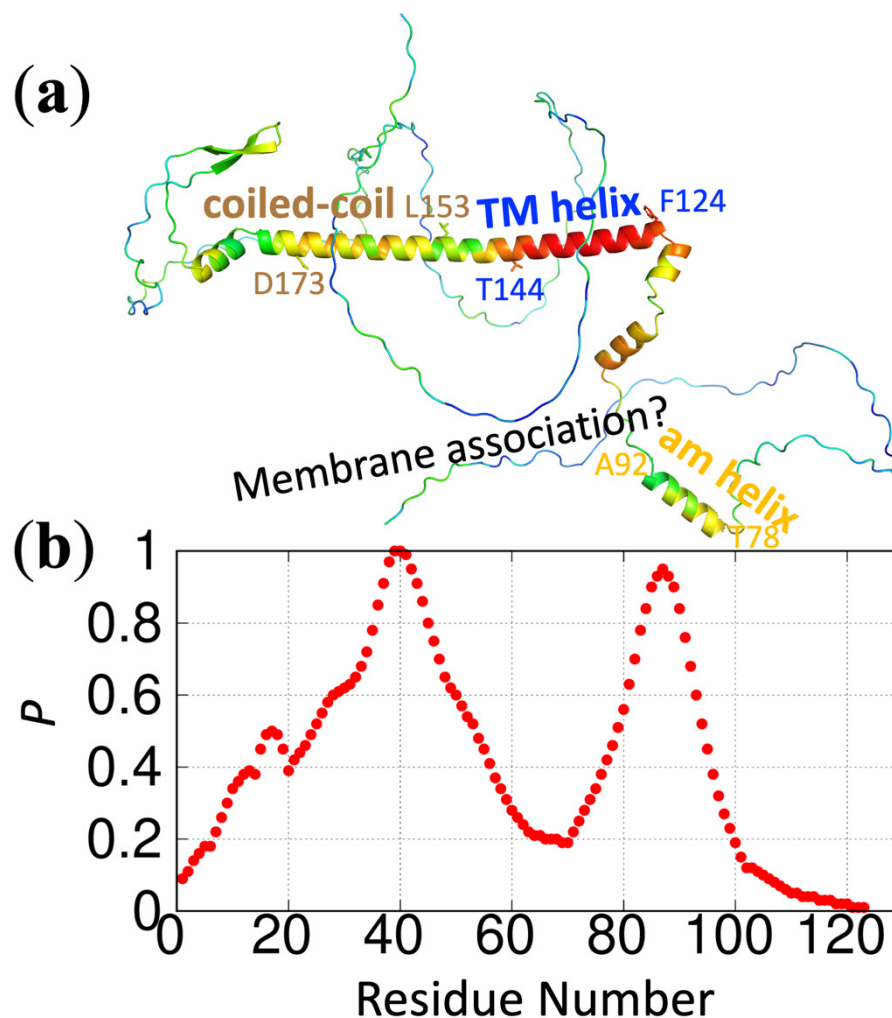


Figure 4. Predicted membrane-association propensities of FtsL. (a) Predicted structure by AlphaFold (<https://alphafold.ebi.ac.uk/>, accessed on 18 March 2022) [18]. The color spectrum displays confidence levels of prediction, ranging from red for high confidence to blue for low confidence. Three putative helices are indicated with start and end residues: amphipathic (am) helix, residues T78-A92; transmembrane (TM) helix, residues F124-T144; and coiled-coil, L153-D173. (b) Predicted membrane-association propensities using the sequence of the first 123 residues, high for residues 25–53 and the putative amphipathic helix. The protein sequence is from Uniprot entry O06213.

4. Discussion

There is growing recognition of the biological importance of IDR-membrane binding. ReSMAP, by predicting residue-specific membrane association propensities of intrinsically ordered proteins, fulfills an urgent need where little prior work has been conducted. Amphipathic helices have received some attention. In particular, the concept of the hydrophobic moment has been introduced [19] and used [20] to analyze known or predicted amphipathic helices. Some methods have been developed to predict membrane-binding amphipathic helices. One such method is AmphipaSeek (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_amphipaseek.html, accessed on 1 July 2022) [21], which is a support-vector-machine classifier trained on 21 proteins with in-plane membrane anchors centered on one (or more) amphipathic helix. AmphipaSeek predicted only nine residues (V3-A11) of α -synuclein as membrane anchors, even though NMR spectroscopy has shown that the first 100 residues of this protein form amphipathic helices that stably associate with membranes [13]. AmphipaSeek also failed to predict any membrane anchors for either FtsQ or SepF, even though both of them are known to have an amphipathic helix

for membrane association (Figures 1d and S5a and Ref. [9]). In comparison, our ReSMAP method predicts high membrane-association propensities for all these amphipathic helices (Figures 1d and S5a,b). Another recent method [22], based on a convolutional neural network trained on 121 membrane proteins, failed to predict any amphipathic helix for α -synuclein, FtsQ, and SepF.

The ReSMAP method currently groups amino acids into only three types: positively charged (K, R, and the *N*-terminus), negatively charged (D, E, and the *C*-terminus), and neutral (all other amino acids). This net-charged based grouping is the main limitation of ReSMAP at present and is necessitated by the limited training data. While the latter predicament highlights an important area for future work, it also raises the concern of whether the training set collected here is representative of the association of IDRs with acidic membranes in general. The validation provided by the experimental data gathered from the literature on six more proteins (Figure 2) goes a long way in allaying this concern.

As more MD simulation data for IDR-membrane association become available, it may be sensible to divide up the neutral amino acids (e.g., polar vs. nonpolar), or even separate R and K. Additionally, lipid compositions vary widely in cell membranes, in experimental studies, and even in our MD simulations (Figure 2 caption and Table S2). In particular, the acidic lipids, including POPS, PIP₂, and POPG, vary from one study to another. By training on data collected from studies with different lipid compositions, ReSMAP relies on common molecular and biophysical properties of most lipids [23], such as net charge (for acidic lipids), polar headgroups, and hydrophobic tails. It has been suggested that membrane protein structure and function have a high degree of tolerance to the change in lipid composition [24]. Still, membrane-association propensities of an IDR may well change subtly as the lipid composition is varied, and such secondary effects may be important biologically but again will require more MD or experimental data for training. Another subtlety is that IDR association will affect the lipid distribution within a membrane. For example, as illustrated in Figure 1a, acidic lipids (POPG) will cluster around positively charged R residues. ReSMAP does not directly account for such lipid redistribution. Rather, it assumes, for a given residue, an average effect on its own membrane-association propensity and on those of its neighbors.

The prediction accuracy of ReSMAP for IDRs with both amphipathic helices and disordered residues is already good but is still lower than the accuracy for fully disordered IDRs. Amphipathic helices and disordered residues likely have different driving forces for membrane association. Amphipathic helices, by definition, have a charged face and a nonpolar face, and are inserted more deeply into the membrane, such that charged groups interact with lipid headgroups while nonpolar groups interact with lipid acyl tails. In contrast, disordered residues are less buried and mainly interact with lipid headgroups. For IDRs containing amphipathic helices, the ReSMAP optimized parameters are a compromise between the helical and disordered residues. A future development will be to increase the number of parameters and train on a data set with more amphipathic helices.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/membranes12080773/s1>, Figure S1: Conversion from z_{tip} to membrane-contact probability; Figure S2: Correlation of MD membrane-contact probabilities with the seven-residue moving average charge for six fully disordered IDRs; Figure S3: Comparison of MD membrane-contact probabilities and predicted membrane-association propensities; Figure S4: Correlation between MD or NMR membrane-contact probabilities and those predicted by ReSMAP; Figure S5: Comparison of MD or NMR membrane-contact probabilities and predicted membrane-association propensities; Table S1: Sequences of eight IDRs; Table S2: Lipid compositions of membranes in the MD simulations; Table S3: Amplitude parameters.

Author Contributions: Conceptualization, S.Q., A.H. and H.-X.Z.; methodology, S.Q. and A.H.; software, S.Q.; data curation, A.H., S.D. and R.P.; writing—original draft preparation, H.-X.Z.; writing—review and editing, H.-X.Z.; supervision, H.-X.Z.; project administration, H.-X.Z.; funding acquisition, H.-X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Institutes of Health, grant numbers GM118091 and AI119178.

Data Availability Statement: The software developed here is accessible and can be downloaded at <https://pipe.rcc.fsu.edu/ReSMAPidp/>. The data used for training the method can be requested from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Kjaergaard, M.; Kragelund, B.B. Functions of intrinsic disorder in transmembrane proteins. *Cell Mol. Life Sci.* **2017**, *74*, 3205–3224. [[CrossRef](#)] [[PubMed](#)]
2. Prehoda, K.E.; Scott, J.A.; Mullins, R.D.; Lim, W.A. Integration of multiple signals through cooperative regulation of the N-WASP-Arp2/3 complex. *Science* **2000**, *290*, 801–806. [[CrossRef](#)] [[PubMed](#)]
3. McLaughlin, S.; Murray, D. Plasma membrane phosphoinositide organization by protein electrostatics. *Nature* **2005**, *438*, 605–611. [[CrossRef](#)]
4. Maki, B.A.; Aman, T.K.; Amico-Ruvio, S.A.; Kussius, C.L.; Popescu, G.K. C-terminal domains of N-methyl-D-aspartic acid receptor modulate unitary channel conductance and gating. *J. Biol. Chem.* **2012**, *287*, 36071–36080. [[CrossRef](#)] [[PubMed](#)]
5. Hicks, A.; Escobar, C.A.; Cross, T.A.; Zhou, H.X. Fuzzy association of an intrinsically disordered protein with acidic membranes. *JACS Au* **2021**, *1*, 66–78. [[CrossRef](#)] [[PubMed](#)]
6. Haxholm, G.W.; Nikolajsen, L.F.; Olsen, J.G.; Fredsted, J.; Larsen, F.H.; Goffin, V.; Pedersen, S.F.; Brooks, A.J.; Waters, M.J.; Kragelund, B.B. Intrinsically disordered cytoplasmic domains of two cytokine receptors mediate conserved interactions with membranes. *Biochem. J.* **2015**, *468*, 495–506. [[CrossRef](#)]
7. Zhang, H.; Cordoba, S.-P.; Dushek, O.; Anton van der Merwe, P. Basic residues in the T-cell receptor ζ cytoplasmic domain mediate membrane association and modulate signaling. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 19323–19328. [[CrossRef](#)] [[PubMed](#)]
8. Sommer, A.; Kordowski, F.; Büch, J.; Marezky, T.; Evers, A.; Andrä, J.; Düsterhöft, S.; Michalek, M.; Lorenzen, I.; Somasundaram, P.; et al. Phosphatidylserine exposure is required for ADAM17 sheddase function. *Nat. Commun.* **2016**, *7*, 11523. [[CrossRef](#)] [[PubMed](#)]
9. Sogues, A.; Martinez, M.; Gaday, Q.; Ben Assaya, M.; Graña, M.; Voegelé, A.; VanNieuwenhze, M.; England, P.; Haouz, A.; Chenal, A.; et al. Essential dynamic interdependence of FtsZ and SepF for Z-ring and septum formation in *Corynebacterium glutamicum*. *Nat. Commun.* **2020**, *11*, 1641. [[CrossRef](#)]
10. Sigal, C.T.; Zhou, W.; Buser, C.A.; McLaughlin, S.; Resh, M.D. Amino-terminal basic residues of Src mediate membrane binding through electrostatic interaction with acidic phospholipids. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12253–12257. [[CrossRef](#)]
11. Pond, M.P.; Eells, R.; Treece, B.W.; Heinrich, F.; Lösche, M.; Roux, B. Membrane Anchoring of Hck Kinase via the Intrinsically Disordered SH4-U and Length Scale Associated with Subcellular Localization. *J. Mol. Biol.* **2020**, *432*, 2985–2997. [[CrossRef](#)] [[PubMed](#)]
12. Lakomek, N.-A.; Yavuz, H.; Jahn, R.; Pérez-Lara, Á. Structural dynamics and transient lipid binding of synaptobrevin-2 tune SNARE assembly and membrane fusion. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 8699–8708. [[CrossRef](#)] [[PubMed](#)]
13. Jacob, R.S.; Eichmann, C.; Dema, A.; Mercadante, D.; Selenko, P. alpha-Synuclein plasma membrane localization correlates with cellular phosphatidylinositol polyphosphate levels. *eLife* **2021**, *10*, e61951. [[CrossRef](#)] [[PubMed](#)]
14. Li, D.W.; Xie, M.; Bruschiweiler, R. Quantitative cooperative binding model for intrinsically disordered proteins interacting with nanomaterials. *J. Am. Chem. Soc.* **2020**, *142*, 10730–10738. [[CrossRef](#)]
15. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [[CrossRef](#)]
16. Piana, S.; Donchev, A.G.; Robustelli, P.; Shaw, D.E. Water dispersion interactions strongly influence simulated structural properties of disordered protein states. *J. Phys. Chem. B* **2015**, *119*, 5113–5123. [[CrossRef](#)]
17. Gould, I.R.; Skjevik, A.A.; Dickson, C.J.; Madej, B.D.; Walker, R.C. Lipid17: A comprehensive AMBER force field for the simulation of zwitterionic and anionic lipids. 2019; *in preparation*.
18. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)]
19. Eisenberg, D.; Schwarz, E.; Komaromy, M.; Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **1984**, *179*, 125–142. [[CrossRef](#)]
20. Gautier, R.; Douguet, D.; Antony, B.; Drin, G. HELIQUEST: A web server to screen sequences with specific α -helical properties. *Bioinform* **2008**, *24*, 2101–2102. [[CrossRef](#)]
21. Sapay, N.; Guermeur, Y.; Deléage, G. Prediction of amphipathic in-plane membrane anchors in monotopic proteins using a SVM classifier. *BMC Bioinform.* **2006**, *7*, 255. [[CrossRef](#)]
22. Feng, S.H.; Xia, C.Q.; Zhang, P.D.; Shen, H.B. Ab-Initio Membrane Protein Amphipathic Helix Structure Prediction Using Deep Neural Networks. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2022**, *19*, 795–805. [[CrossRef](#)] [[PubMed](#)]

-
23. Zhou, H.-X.; Cross, T.A. Influences of Membrane Mimetic Environments on Membrane Protein Structures. *Annu. Rev. Biophys.* **2013**, *42*, 361–392. [[CrossRef](#)] [[PubMed](#)]
 24. Sanders, C.R.; Mittendorf, K.F. Tolerance to Changes in Membrane Lipid Composition as a Selected Trait of Membrane Proteins. *Biochemistry* **2011**, *50*, 7858–7867. [[CrossRef](#)] [[PubMed](#)]