Quantification of Conformational Entropy Unravels Effect of Disordered Flanking Region in Coupled Folding and Binding

Theisen, Frederik Friis; Staby, Lasse; Tidemand, Frederik Grønbæk; O’Shea, Charlotte; Prestel, Andreas; Willemoës, Martin; Kragelund, Birthe B.; Skriver, Karen

Published in:
Journal of the American Chemical Society

DOI:
10.1021/jacs.1c04214

Publication date:
2021

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):
Quantification of Conformational Entropy Unravels Effect of Disordered Flanking Region in Coupled Folding and Binding

Frederik Friis Theisen, Lasse Staby, Frederik Grønbæk Tidemand, Charlotte O’Shea, Andreas Prestel, Martin Willempøës, Birthe B. Kragelund,* and Karen Skriver*

ABSTRACT: Intrinsic disorder (ID) constitutes a new dimension to the protein structure–function relationship. The ability to undergo conformational changes upon binding is a key property of intrinsically disordered proteins and remains challenging to study using conventional methods. A 1994 paper by R. S. Spolar and M. T. Record presented a thermodynamic approach for estimating changes in conformational entropy based on heat capacity changes, allowing quantification of residues folding upon binding. Here, we adapt the method for studies of intrinsically disordered proteins. We integrate additional data to provide a broader experimental foundation for the underlying relations and, based on >500 protein–protein complexes involving disordered proteins, reassess a key relation between polar and nonpolar surface area changes, previously determined using globular protein folding. We demonstrate the improved suitability of the adapted method to studies of the folded αα-hub domain RST from radical-induced cell death 1, whose interactome is characterized by ID. From extensive thermodynamic data, quantifying the conformational entropy changes upon binding, and comparison to the NMR structure, the adapted method improves accuracy for ID-based studies. Moreover, we apply the method, in conjunction with NMR, to reveal hitherto undetected effects of interaction–motif context. Thus, inclusion of the disordered context of the DREB2A RST-binding motif induces structuring of the binding motif, resulting in major enthalpy–entropy compensation in the interaction interface. This study, also evaluating additional interactions, demonstrates the strength of the ID-adapted Spolar–Record thermodynamic approach for dissection of structural features of ID-based interactions, easily overlooked in traditional studies, and for translation of these into mechanistic knowledge.

INTRODUCTION

The structure–function paradigm of proteins has recently been challenged by the emergence of intrinsic disorder (ID) as a functional property. In contrast to globular protein domains, intrinsically disordered regions (IDRs) are not adequately characterized by a single structure. Instead, IDRs populate conformational ensembles.1−3 This property likely arises because IDRs, compared to globular proteins, are enriched in polar and charged residues and depleted in hydrophobic ones.4−9

Since IDRs do not adopt stable structures in the unbound state, their interactions are often mediated by sequence motifs, such as short linear motifs (SLiMs). SLiMs are short sequence stretches defined from the primary structure which, when present in IDRs, may form secondary structure in complex.7 Though ID interactions often involve structuring of the disordered protein through coupled folding and binding,6 the resulting complex may retain considerable dynamics, as seen for the transcriptional regulator Gcn4, where an α-helix binds in multiple orientations.9

While concepts such as fuzziness10 may reduce the loss of conformational entropy associated with IDR binding, any ordering of the interacting proteins will result in decreased conformational entropy. This entropic penalty is a frequently discussed property of ID-based interactions.11 However, since the total interaction entropy contains several underlying terms, it is often challenging to study the isolated conformational contribution experimentally.

A paper by R. S. Spolar and M. T. Record, published in 1994, presented a method for quantifying the loss of conformational entropy associated with protein interactions using an empirically derived thermodynamic relation and correlated this to the number of residues folding upon binding, $R_{\Delta S}$.12 The method has previously been used to quantify...
coupled folding and binding of several protein–protein and protein–ligand interactions.\(^{12–14}\) Generally, parametrization of protein–protein interaction thermodynamics allows the entropic contribution to be expressed as follows: \(\Delta S^0(T) = \Delta S^0_{\text{hydration}}(T) + \Delta S^0_{\text{conformation}} + \Delta S^0_{\text{interactions}}\) where the \(\Delta S^0_{\text{interactions}}\) term describes rotational and translational contributions to entropy and is proposed to be a constant, interaction-independent value.\(^{12,15}\) In their 1994 paper, Spolar and Record also proposed an empirical relation between heat capacity change upon binding, which can be experimentally determined, and hydration entropy. This relies on three main assumptions: (1) the heat capacity change of protein interactions is dominated by dehydration of polar and nonpolar surfaces, (2) hydration entropy of protein interactions is dominated by nonpolar surface dehydration, and (3) for protein binding and folding the ratio of polar to nonpolar surface area burial is approximately constant.\(^{12}\) These assumptions enable an empiric approxmation of the hydration entropy as a function of heat capacity change through the following equation: \(\Delta S^0_{\text{hydration}}(T) = c_{\text{nonpolar}} \Delta A S_{\text{nonpolar}} \ln(T/386 \text{ K})\), where \(c_{\text{nonpolar}}\) is the heat capacity coefficient for hydration of nonpolar protein surface and 386 K reflects observations that hydrophobic hydration entropy extrapolates to zero at 386 K.\(^{16}\) Within the relevant temperature range, the entropic contribution from hydration of polar surface was previously found not to be significant\(^{12}\) and was therefore also disregarded in this work.

While the 1994 paper provided a means for dissecting interaction entropy using thermodynamic data, the method was developed prior to the recognition of ID as a functional property. In addition, subsequent determination of the parameters conflicted with several of the originally assumed relations (Tables 1 and 2).\(^{15,17}\) These important limitations jointly prompted us to reevaluate the method and underlying parameters.

<table>
<thead>
<tr>
<th>source</th>
<th>data set</th>
<th>(c_{\text{nonpolar}}^{b})</th>
<th>(c_{\text{polar}}^{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spolar et al.(^{2,11})</td>
<td>amides + hydrocarbons</td>
<td>1.3 ± 0.2</td>
<td>−0.6 ± 0.2</td>
</tr>
<tr>
<td>Myers et al.(^{12})</td>
<td>26 proteins</td>
<td>1.2 ± 0.5</td>
<td>−0.4 ± 1.3</td>
</tr>
<tr>
<td>Makhatadze and Privalov(^{13})</td>
<td>20 proteins</td>
<td>2.1 ± 0.4</td>
<td>−0.9 ± 0.8</td>
</tr>
<tr>
<td>Robertson and Murphy(^{14})</td>
<td>49 proteins</td>
<td>0.7 ± 0.2</td>
<td>0.5 ± 0.3</td>
</tr>
<tr>
<td>Murphy and Freire(^{14})</td>
<td>cyclic dipeptides</td>
<td>1.88 ± 0.08</td>
<td>−1.1 ± 0.1</td>
</tr>
<tr>
<td>this work(^{c})</td>
<td></td>
<td>1.37 ± 0.06</td>
<td>−0.6 ± 0.1</td>
</tr>
</tbody>
</table>

"This work incorporates data derived from both protein unfolding and from hydrocarbon and cyclic dipeptide solubility to determine the surface hydration coefficients using a single two-variable linear fit (Figure S1A).\(^{c}\) Coefficients given in \(\text{J mol}^{-1} \text{K}^{-1} \text{Å}^{-2}\) and relevant in the empiric relation \(\Delta C_p = c_{\text{nonpolar}} \Delta A S_{\text{nonpolar}} + c_{\text{polar}} \Delta A S_{\text{polar}}\).\(^{c}\) Full list of entries is shown in Table S1.

In this study, we revisit the Spolar–Record (SR) method in an ID context and demonstrate how in-depth analysis of thermodynamic parameters may provide an enhanced understanding of the structural transitions of protein–protein interactions. To this end, we use the plant-based radical-induced cell death 1 (RCD1):transcription factor (TF) ID interaction system.\(^{18}\) The model involves the small folded αα-hub domain RST (RCD1-SRO1-TAF4) from RCD1,\(^{19,20}\) which mediates interactions with a large number of disordered partners.\(^{21,22}\) As an αα-hairpin, the RST domain has a central αα-hairpin super secondary structure consisting of two antiparallel α-helices connecting two additional α-helices. The αα-hairpin is rigid and forms a stable platform that organizes more dynamic α-helices, allowing adaptation to different binding partners.\(^{19,23}\) Dehydration responsive element binding-protein 2A (DREB2A), a TF involved in abiotic plant stress responses,\(^{24}\) is a well-characterized interaction partner of RCD1-RST and interacts with the RST domain through a SLiM found in its long C-terminal IDR, which folds upon binding to form a short α-helix in the complex.\(^{19,22}\) While the interaction of the SLiM is well-characterized, studies that include the flanking regions suggest that these also affect the interaction,\(^{22,25}\) in accordance with the increasing appreciation of context contributions in SLiM-based interactions.\(^{26}\) With an outset in the ID-adapted SR method we dissect the entropic contribution of the RCD1-binding SLiM flanking regions and employ NMR spectroscopy to probe the origin of conformational entropy changes.

**RESULTS AND DISCUSSION**

Adapting the SR Method for Use with ID-Based Interactions. When we considered the SR method for in-depth analysis of ID-related binding thermodynamics, we evaluated recently published literature to provide a more robust experimental foundation for the empiric relations key to this method. In particular, the coefficients (\(c_{\text{polar}}\) and \(c_{\text{nonpolar}}\)) relating surface area changes to heat capacity changes have been subject to considerable variations when determined by different research groups (Table 1) (reviewed in Prabhu and Sharp, 2005).\(^{27}\) While the majority agree on the sign of the contributions, with solvent exposure of nonpolar surface resulting in a large increase in heat capacity, the specific coefficients vary considerably depending on the data set.

Due to the covariance of polar and nonpolar surface area changes in the folding of globular proteins,\(^{28,29}\) the two cannot be considered independent and thus do not provide sufficiently strong constraints for reliable fitting. This issue has been mitigated by using hydrocarbon compounds to determine the nonpolar coefficient separately.\(^{29}\) In addition, dipeptides have been used as model compounds to determine the coefficients. Here, we combined the three data types and determined both coefficients using a single two-variable fit (Figure S1 and Table 1). Interestingly, this expanded data set produced coefficient values similar to those originally used by Spolar and Record (SR).\(^{12}\) The validity of using coefficients derived from globular protein folding is central, and recent work indicates that ID-based interactions are driven by the same hydrophobic effect...
that govern globular protein folding,\textsuperscript{30} thus warranting the application of the same coefficients for both protein folding and ID-based interactions.

While dehydration of protein surface may be comparable for globular proteins and ID proteins, these proteins differ in their most fundamental property: their amino acid composition.\textsuperscript{35} Since the SR analysis exploits an apparent ratio between changes in polar and nonpolar surface area, it is important that this ratio is applicable in an ID context. The original ratio is based on globular protein folding, a process that is largely driven by favorable desolvation in a so-called “hydrophobic collapse”, which may not represent ID-based protein–protein interactions. To determine an appropriate ratio for ID-based interactions, we analyzed data from the Disordered Binding Site (DIBS) database (Table S2).\textsuperscript{36} The database contains protein complexes involving at least one disorderd partner. Complex structures were filtered as described in Methods, resulting in 577 suitable entries. The difference in polar and nonpolar surface area between the complex and the two individual proteins on their own was calculated (Figure 1), and

\begin{equation}
Δ\text{ASA}_{\text{polar}} = (0.92 \pm 0.03)Δ\text{ASA}_{\text{nonpolar}}.
\end{equation}

Figure 1. Relation between changes in polar and nonpolar surface area for ID-based protein–protein interactions. Each data point represents the $Δ\text{ASA}$ between the complex and unbound states for a protein–protein interaction involving one intrinsically disordered interaction partner. Data collection and processing are described in the Methods. The data were fitted using simple linear regression, and the solid line represents the weighted linear fit with a slope of $0.92 \pm 0.03$. The dashed line represents the ratio used in the original SR method.\textsuperscript{12,29}

the distribution of ID chain lengths is shown in Figure S8. Since the length of the ID component was generally less than 30 residues, the unbound state was modeled using random-coil properties. Weighted linear fitting of the 577 protein complexes resulted in the following linear correlation: $Δ\text{ASA}_{\text{polar}} = (0.92 \pm 0.03)Δ\text{ASA}_{\text{nonpolar}}$. This ratio is significantly higher than the 0.59 used in the original SR analysis\textsuperscript{12,29} and will impact analysis of ID-based interactions. It should be noted that the crystal structures only define the comparatively rigid regions of a protein complex, leaving dynamic regions unrepresented in a PDB file. The distribution of ID component lengths (Figure S8) suggests a binding site size range for which the ID adaption is most appropriate.

The rotational and translational entropy change was originally estimated using four “rigid body” enzyme–ligand interactions. In addition to being a relatively small number of data points, the emerging view is that enzymes have significant internal motion, which is influenced by substrate binding,\textsuperscript{57} and consequently their interactions do not represent simple rigid body docking events. Thus, to accommodate this, we looked at published results based on both theoretical and experimental evaluations as well as molecular dynamics simulations (Table 2). In agreement with Spolar and Record, a study investigating 28 protein–ligand interactions concluded that the rotational and translational entropy change is constant and independent of the specific interaction.\textsuperscript{15} This conclusion was corroborated by molecular dynamics simulations, which suggested similar and constant values of $ΔS^0_{\text{rot}}$ for macro-molecular interactions.\textsuperscript{38,39} Based on these observations, an interaction-independent $ΔS^0_{\text{rot}}$ value of $-110$ J mol$^{-1}$ K$^{-1}$, differing substantially from the original value, was used in this work. Reevaluation of a multimeric complex assembly, which has larger rotational and translational contributions, also analyzed by Spolar and Record, suggests that the reduced $ΔS^0_{\text{rot}}$ value improves that accuracy of the $R_\text{m}$ value (Table S3).

To summarize, we introduced additional data to provide a more robust experimental foundation (Tables 1 and 2). In addition, we revisited surface area changes in the context of ID-based interactions and adjusted according to our findings (Figure 1). This resulted in the following modified equation (originally derived by R. S. Spolar and M. T. Record)\textsuperscript{12} for the ID-adapted SR method ($SR_{\text{id}}$):

\begin{equation}
ΔS^0_{\text{conformational}} = -1.66ΔC_p \ln(T/T_0) + 110 \ J \text{ mol}^{-1} \text{ K}^{-1}
\end{equation}

where $ΔC_p$ is the experimentally determined heat capacity change upon binding and $T_0$ is the isentropic ($ΔS^0(T_0) = 0$) temperature. These two variables can be determined using isothermal titration calorimetry (ITC) by conducting experiments across a range of temperatures. Thermodynamic quantification of the number of residues folding upon binding, $R_{\text{id}}$, can then be calculated by dividing $ΔS^0_{\text{conformational}}$ by the average per-residue loss of conformational entropy for protein folding.\textsuperscript{32} Since the expanded data set led to a small adjustment of the surface hydration coefficients (Table 1), the average per-residue loss of conformational entropy was adjusted to $-24.0$ J mol$^{-1}$ K$^{-1}$ residue$^{-1}$ (originally $-23.4$ J mol$^{-1}$ K$^{-1}$ residue$^{-1}$). This value is based on globular protein folding; however, adjusting for the compositional bias of ID regions did not result in significant changes (Methods). With the adapted $SR_{\text{id}}$ equation, we confirmed the method using a well-studied ID-based protein interaction. For this, we revisited the RCD1-RST interaction with DREB2A, which provided both a target for verification and also opportunities for further analysis.

RCD1-RST Interacting Region in DREB2A. Previous studies revealed that inclusion of regions flanking the RCD1-binding SLiM of DREB2A increased the affinity for RCD1-RST.\textsuperscript{22,25} To study how the DREB2A SLiM flanking regions influence the interaction with RCD1-RST, we determined the thermodynamic parameters of several different DREB2A fragments using ITC (Figure 2, Figure 3, and Table 3). Results for each fragment were derived from more than 10 experiments conducted over a range of temperatures.

Examining the affinities, only the N-terminal extension from the SLiM affected the affinity significantly, with DREB2A$_{243-272}$ and DREB2A$_{244-287}$ displaying the strongest binding, both in the low nM regime. While the specific thermodynamic properties of the two high-affinity fragments differed slightly, the differences were too small to argue that residues outside DREB2A$_{243-272}$ influenced the interaction.
derived from the temperature dependence shown in Figure 4, and reported errors are the propagated standard deviations of the linear coefficients calculated as described in the Methods. The $K_d$ values are calculated from the Gibbs free energy change at 25 °C.

Further analysis of the thermodynamic parameters (Table 3) revealed that the increase in affinity for the high-affinity DREB2A fragments was driven by enthalpic contributions. Interestingly, the enthalpy change per residue of the longer DREB2A255–272 fragment was roughly equal to that of the longer DREB2A243–272 fragment. However, the longer DREB2A fragment was distinguished by a large unfavorable entropic contribution, which offsets most of the favorable enthalpy. This further supported the entire 243–272 region of DREB2A being involved in RCD1 binding.

Conformational Entropy Penalty of N-Terminal Flanking Region. The entropy change of the longer DREB2A fragments is intriguing, as structural predictions (Figure 2) and nuclear magnetic resonance (NMR) studies of the short DREB2A fragment did not support formation of secondary structure outside the established $\alpha$-helix, spanning residues V261–G270. To develop a better understanding of how the flanking regions contribute to the interaction, we employed the SRiD method. To determine the $\Delta C_p$ upon binding, we exploited the temperature dependence of the enthalpy change by performing ITC experiments over a range of temperatures (Figure 4). The derived parameters and calculated $R_0$ values are shown in Table 4. $R_0$ values determined using the original ($R_{0n}$) and the ID-adapted ($R_{0ID}$) SR method differ. The higher estimates proposed by the adapted method fit better with previously shown structuring of the short 18-residue DREB2A255–272 fragment in addition to considerable stabilization of the RST domain.19 For ID interactions with RCD1-RST, the original SR method underestimates the conformational entropy loss.

The analysis suggests that the interactions of the N-terminally extended DREB2A fragments result in a higher degree of structuring in the complex. Interestingly, the number of residues estimated to fold upon binding is close to the number of residues from the RCD1-binding region present in each fragment (Figure 2, hatched region). This corroborates the proposed continuous binding region, although it poses the question of why the SLiM region has a net entropy close to zero at 25 °C while inclusion of the N-terminal flanking region results in a considerable unfavorable total entropy. The unfavorable total entropy change of DREB2A243–272 binding suggests that the increased structuring does not produce comparable favorable desolvation.

DREB2A N-Terminal Flanking Region Retains Disorder in Complex. Driven by the unexpected thermodynamics of the flanking regions, we set out to study the
interaction of the RCD1-binding region using NMR spectroscopy. Previous studies of the RCD1-RST:DREB2A complex using the short DREB2A235–276 fragment were somewhat impeded by missing peaks from residues of both RCD1-RST and DREB2A. To monitor changes in the flanking regions, we introduced an extended DREB2A fragment spanning residues Q234–G276. Using this DREB2A234–276 fragment resulted in generally improved spectral quality of the bound state of both DREB2A and RCD1-RST. Consequently, we obtained almost complete backbone assignments of both interaction partners in the bound state and of DREB2A in the free state.

The interaction was studied by monitoring $^{15}$N,HN chemical shift perturbations (CSPs), secondary chemical shifts, and $^{13}$C and $^1$H chemical shifts of free and bound DREB2A. Values for the unbound state are shown in Figure S2. Representative isotherms are shown in Figures S2–S5. Error bars represent standard deviations of two or more experiments. Standard deviations smaller than 4 kJ mol$^{-1}$ are not drawn.

**Table 4. Thermodynamic Estimates of the Number of Residues Undergoing Coupled Folding and Binding for the RCD1:DREB2A Interaction**

<table>
<thead>
<tr>
<th>DREB2A</th>
<th>$T_s$ (K)</th>
<th>$\Delta C_p$ (J mol$^{-1}$ K$^{-1}$)</th>
<th>$R_{\Delta G}$</th>
<th>$R_{\Delta H}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>255–272</td>
<td>295.3 ± 1.1</td>
<td>−1.3 ± 0.2</td>
<td>11 ± 4</td>
<td>20 ± 5</td>
</tr>
<tr>
<td>250–287</td>
<td>297.2 ± 0.7</td>
<td>−1.4 ± 0.1</td>
<td>13 ± 2</td>
<td>22 ± 2</td>
</tr>
<tr>
<td>243–272</td>
<td>286.4 ± 1.0</td>
<td>−1.8 ± 0.1</td>
<td>22 ± 2</td>
<td>32 ± 3</td>
</tr>
<tr>
<td>234–287</td>
<td>286.9 ± 0.6</td>
<td>−1.8 ± 0.1</td>
<td>21 ± 2</td>
<td>31 ± 2</td>
</tr>
</tbody>
</table>

"Conformational entropy loss quantified as number of residues folding upon binding ($R_{\Delta G}$). $T_s$ is the isentropic temperature. ID refers to the SRID method developed here.

were dominated by residues D257–G270, which are the residues that are buried in the binding groove of RCD1-RST. In contrast, residues in the N-terminal flanking region displayed substantially smaller CSPs. $^{13}$C secondary chemical shifts of free and bound DREB2A verified the previously reported induction of $\alpha$-helical structure upon binding to RCD1-RST.

In addition, secondary chemical shifts of residues D257–D260 indicated formation of previously unidentified, highly populated extended structure in the complex. Apart from that, the N-terminal flanking region did not form additional secondary structure (Figure 5C). Analysis of chemical shifts of the bound state of DREB2A using the motif identification from chemical shifts (MICS) server$^{62}$ returned a high probability that residue D260 functions as an Ncap of the V261–G270 $\alpha$-helix (Figure S6F). This role of D260 as an
These clusters were located in continuation of the N-terminus elevated flod dynamics of the DREB2A N-terminal DREB2A majority of RCD1-RST CSPs induced by inclusion of the DREB2A fragment stabilizes existing RST structure. While the di allow almost complete assignment of the RST domain. Still, larger CSPs were seen in helix 1 and 2 (Figures 6E and 7B). (Figure 6D and E, Figure S7A long and the short DREB2A fragments were relatively small which are present in DREB2A243 discrete interactions of residues G243, D247, and S251, all of which are present in DREB2A245−272. Although the CSPs and secondary shifts of the N-terminal flanking region are relatively small (<0.07 ppm, Figure 5B) and provided no evidence of substantial structuring (Figure S6G), the conformational dynamics may be affected, which would potentially explain the loss of conformational entropy (Table 4).

To study how binding to RCD1-RST influences the dynamics of the DREB2A N-terminal flanking region, we acquired 1H relaxation rates (Figures 5D and S6H and I) of DREB2A. Relaxation rates of the RCD1-binding SLiM region were strongly affected by complex formation and closely resembled those obtained for the bound RCD1-RST (Figure 5D, Figure S6H and I, dashed lines), indicating tumbling comparable to the folded RST domain. Interestingly, the N-terminal flanking region showed comparatively small changes; however, the effect extended further than the persistence length of a disordered protein, indicating a partly restricted conformational space. This may suggest that the A240-D254 region participates in transient interactions with RCD1-RST, as indicated by the elevated Rs (Figure S6). While the NMR analysis revealed interesting properties of DREB2A in complex, it also indicated that the loss of conformational entropy cannot be attributed solely to structuring of DREB2A.

**Enthalpy—Entropy Compensation Driven by Flanking Disorder in Coupled Folding and Binding.** NMR spectra of labeled RCD1-RST499−573 in complex with DREB2A234−276 produced considerably improved peak intensities, compared to the complex with DREB2A255−272 (Figure 6A, B, and C), allowing almost complete assignment of the RST domain. Still, the differences in RCD1-RST chemical shifts comparing the long and the short DREB2A fragments were relatively small (Figure 6D and E, Figure S7A−E), suggesting that the longer DREB2A fragment stabilizes existing RST structure. While the majority of RCD1-RST CSPs induced by inclusion of the DREB2A flanking regions are negligible, two small clusters of larger CSPs were seen in helix 1 and 2 (Figures 6E and 7B). These clusters were located in continuation of the N-terminus of DREB2A255−272 in the complex, which suggested interaction between the extended DREB2A N-terminus and helix 1 and 2 of the RST domain. The new assignment of the RCD1-RST N-terminus revealed a few residues, W507, M508, and P509, with secondary chemical shifts indicating extended structure (Figure S7D). The similar extended structure seen in bound DREB2A may thus likely be spatially close to the potential β-strand residues in RCD1-RST, indicating formation of a short intermolecular β-sheet structure. Comparison of RCD1-RST 1H relaxation rates, measured in complex with DREB2A234−276 and DREB2A255−272, showed reduced R1 and elevated R2 rates for the longer DREB2A fragment (Figure S7F and G). The effect was similar for most residues, which is consistent with a slower global tumbling rate resulting from binding of the longer DREB2A234−276 fragment.

Changes to RCD1-RST dynamics upon binding of the short DREB2A234−276 fragment (Figure S7F and G) suggested that considerable conformational entropy was lost to stabilization of the RST structure. Given that the conformational entropy loss (Table 4, R0=1.8) applies to the entire system, DREB2A234−276 is not fully ordered in complex. This is corroborated by published NMR results, which suggested that DREB2A dynamics persist in the RCD1-RST:DREB2A234−276 complex. Thus, it seems plausible that inclusion of the DREB2A N-terminal flanking region of the RCD1-binding SLiM stabilizes existing but dynamic, complex structure (Figure 7), resulting in more favorable interactions but also a considerable loss of conformational entropy. In particular, the appearance of NMR signals belonging to the Ncap region of the DREB2A bound state α-helix is interesting, since the region is proposed to form a hydrophobic staple motif. Stabilization of this motif is thus likely to have an effect on helix stability. In addition, the staple motif is likely to be located near helix 1 of the RCD1-RST domain. As a result, stabilization of the staple motif may cooperatively stabilize helix 1 of the RST domain, explaining the improved RST NMR peak intensities in complex with DREB2A255−272 and the slightly elevated secondary chemical shifts (Figure S7C). This interaction model implies enthalpy—entropy compensation driven by flanking, high-entropy, interactions.

**Application of the SRID Method on Other ID-Based Interaction Systems.** In this study, we applied the SRID method to the RCD1-RST:DREB2A interaction system, which involves dynamics of mechanistic importance beyond coupled
folding and binding of the core SiLM. Originally, the SR method was evaluated using many different systems ranging from small-molecule interactions to multimeric complex assembly. To test the applicability of the ID-adapted SR method, we applied it to data analyzed both in the original SR paper and in more recent work (Tables S3 and S4). From Spolar and Record, we focused on protein–protein interactions involving prototypical coupled folding and binding of ID regions such as the angiotensin II:antibody Fab 131 and the S-peptide:S-protein interactions. In the first case, the $R_{\text{th, ID}}$ value matched the structure-derived $R_{\text{th}}$ value better than the $R_{\text{th}}$ value did. In the second case, $R_{\text{th, ID}}$ suggested more extensive folding and stabilization which has to include effects from both interaction partners. Recent molecular simulations supported this, as while the complex remained close to the X-ray structure, the isolated S-protein showed increased structural fluctuations across the protein. Literature mining revealed several additional sets of data, where the SR method could be employed to improve the understanding of the interaction system (Table S4 and Figure S8). For example, SR analysis was applied to the cell cycle regulating p27:Cdk2:Cyclin A complex, which involves extensive coupled

Most of the above-mentioned interactions involve large conformational changes and intricate interactions. We therefore also included systems allowing more direct comparison of the SR and SR$_{\text{ID}}$ analysis. The foldable PDZ1 domain of MAGI-1, functioning in cell–cell junctions, is targeted by the E6 oncoprotein from human papilloma virus. Based on NMR analyses, the PDZ1:E6 peptide interaction resulted in extensive structuring of the PDZ1 C-terminal extension as reflected in the SR analyses (Table S4). SR$_{\text{ID}}$ analysis correctly predicted that a mutant PDZ1 domain (PDZ1 SLVGGG) undergoes almost no structuring upon E6 binding, while the $R_{\text{th}}$ value determined using the original SR method is negative, suggesting unfolding, which seems unlikely based on relaxation rates. Structural and thermodynamic data for the interaction between the disordered tail of the TFIIF-associated CTD phosphatase (FCP1) and the winged-helix domain of the Rap74 subunit of TFIIF also allowed comparison of $R_{\text{th}}$ values. With extensive folding of FCP1 upon binding to Rap74, the $R_{\text{th, ID}}$ value of 23 more likely than the $R_{\text{th}}$ value of 1. This is also the case for the interaction between the nuclear localization signal of RelA and the inhibitor IκBα. Here, from NMR data, a folding of ~24 residues in the RelA peptide alone, not counting potential structural changes in IκBα, makes the $R_{\text{th, ID}}$ value of 23 more likely than the $R_{\text{th}}$ value of 14 (Table S4). Based on SR analysis of the RCD1-RST:DREB2A interaction and additional ID-based interactions in this paper, the adapted SR$_{\text{ID}}$ method shows promise for more in-depth explorative entropy decomposition and mechanistic insight.

A potential application of the SR approach could be analysis of liquid–liquid phase separation (LLPS), now emerging as important to biological regulation. ITC has previously been used to study LLPS for polymers and proteins, but not to determine conformational entropy. However, recent studies found distinct conformational entropies of the different phases, suggesting that application of the SR method to LLPS could constitute a future, but promising goal.

CONCLUSION

In this study we explored how thermodynamic data can improve our understanding of ID-based interactions. Revisiting the empirical method proposed in 1994 by Spolar and Record, we implemented adjustments aiming to adapt the method for studying ID-based interactions. Evaluating our results using the RCD1-RST:DREB2A model, as well as additional ID-based protein–protein interaction systems, suggested that the ID adaption improved the accuracy of the SR analysis of ID-based interactions. In-depth analysis of the RCD1-RST:DREB2A interaction revealed a correlation between binding region length and the estimated number of residues folding upon binding. Structural data indicated that this could not be attributed solely to structuring of the disordered interaction partner but included stabilization across the complex. Despite the apparent complexity, we propose that the higher conformational entropy loss, suggested by the ID-adapted
SR_{ID} method, is more accurate for interactions involving prototypical ID proteins.

We applied the SR_{ID} method to analyze the effect of the disordered flanking regions on the RCD1-RST:DREB2A interaction, which revealed a large unfavorable conformational entropy loss, not compensated by favorable desolvation entropy. This was in contrast to the SLiM region, for which the conformational entropy loss was mostly compensated by desolvation. Based on NMR data describing structural changes and dynamics, the conformational entropy loss was unlikely to originate from folding of the flanking regions. Rather, a more likely explanation of the specific thermodynamic parameters is enthalpy—entropy compensation resulting from stabilization of the RST:DREB2A complex structure. Development of this model was driven by quantification of conformational entropy changes revealing significant losses arising from inclusion of residues which remained dynamic in the bound state. This highlights how the method can provide important insight into additional effects imposed by binding, which are dissipated throughout the complex. NMR spectroscopy can subsequently be employed to unravel where and how these effects are manifested.

From a functional point of view, the influence of the flanking region on the affinity is interesting. While the SLiM is highly conserved within DREB2A orthologs and thus evolutionary constrained, the N-terminal flanking region displays poor conservation. The limited sequence constraints of the flanking regions may thus enable fine-tuning of interaction affinity and specificity. These observations open new possibilities for studying enthalpy—entropy compensation in an evolutionary context using the established RCD1:TF model system.

### METHODS

**Analysis of Solvation Heat Capacity Changes.** In order to circumvent the covariation of protein data, resulting from the apparent correlation of polar and nonpolar surface area changes for protein folding, independent data in the form of hydrocarbon and cyclic dipeptide solvation data were incorporated. Due to the difference in size between small molecules and proteins, normalization was necessary. This was done within each data group by scaling all variables by the maximum heat capacity change observed for that group. Heat capacity coefficients for solvation of polar and nonpolar surface were then determined using a two-variable linear fit. The final data set includes 80 protein entries, 13 hydrocarbon compounds, and six cyclic dipeptides (Table S1).

**Calculation of Disordered Protein Binding Surface Area Changes.** Structures of protein—protein complexes involving disordered protein regions were acquired from the DIBS database (version: 28-09-2017, source: dibs.enzim.ttk.mta.hu). Entries were filtered by specifying a minimum ID chain length of four residues and removal entries with modified residues in the ID chain, as these could not be processed correctly. Entries producing significantly diverging results were manually inspected. The unfiltered list of entries is given in Table S2.

The accessible surface area of complexes and folded components alone was determined using the PyMOL “get_area” command. A surface originating from nitrogen and oxygen atoms was defined as polar, while a surface originating from carbon and sulfur atoms was defined as nonpolar. Hydrogen atoms were not considered explicitly. The implementation is shown in Script 1 and is available from GitHub. Surface area of the disordered component in the unbound state was calculated using a library of amino acid random coil surface areas. This was used along with random coil hydrophobicity percentage to estimate polar and nonpolar surface area of the unbound ID protein.

Surface area changes were calculated by taking the sum of the folded and disordered component in their unbound state and subtracting the surface area of the complex. Each entry was weighted based on ID content prediction by the IUPred2 short disorder tool. The IUPred2 weight was applied using the sigmoidal function: weight = 1 / (1 + e^{−(sID − 1)/(1 + e^{−(sID − 1)})}), where sID is the IUPred2-predicted disorder content of the ID chain. The surface area ratio was extracted from the slope of a linear fit of polar surface area change as a function of nonpolar surface area change (Figure 1).

**Recalibration of the per-Residue Conformational Entropy Loss of Folding.** The loss of entropy per-residue folding was recalculated using the protein folding data referenced by Spolar and Record in Table 1 of R. S. Spolar and M. T. Record 1994, with the slightly adjusted nonpolar surface hydration heat capacity coefficient (Table 1). Effects of amino acid composition differences on the per-residue loss were investigated by calculating the frequency-weighted average folding entropy based on residue-specific values provided by Towse et al.8 Globular and ID protein amino acid frequencies were based on the PDB and DIBS databases, respectively. A resulting difference of less than 1% was deemed insignificant in context of the SR method.

**Protein Production and Purification.** Puriﬁcation of RCD1-RST:DREB2A was performed as previously described. Production and puriﬁcation of DREB2A fragments were done using a glutathione S-transferase (GST) tag and reversed-phase chromatography as described. The GST tag was removed using tobacco etch virus (TEV) protease leaving an N-terminal glycine residue. For NMR, pure protein was transferred to the experimental buffer using size exclusion chromatography.

**Isothermal Titration Calorimetry.** ITC was used to determine thermodynamic parameters of the RCD1-RST:DREB2A interaction. The enthalpy change, ΔH, and affinity, K_0, were determined from the raw ITC data by fitting to the “one set of sites” binding model from the Origin 7 software package. Gbbs free energy, ΔG, and entropy, ΔS, were derived from the affinity and enthalpy. Experiments were performed using a MicroCal ITC-200 microcalorimeter (GE Healthcare). RCD1-RST protein samples were dialyzed against 50 mM HEPES, pH 7.4, and 100 mM NaCl buffer, and freeze-dried DREB2A peptides were resuspended in the filtered dialysis buffer. Samples were degassed either by centrifugation at 20000 g or by vacuum stirring at experimental temperature for 15 min. Thermodynamic parameters at 25 °C were derived from a linear fit of at least 10 experiments conducted at various temperatures using 7–12 μM DREB2A in the cell and 50–100 μM RCD1-RST in the syringe. The number of injections and injection volume varied, depending on interaction affinity, from 19 × 2 μL to 35 × 1 μL with a 180 s equilibration time between injections. Stirring speed was set to 750 rpm. Dilution heats were estimated from the saturation baseline enthalpy and subtracted from all data points.

Heat capacity changes were determined using a series of ITC experiments across a range of temperatures (approximately 10–33 °C) by a linear fit of the temperature dependence of the enthalpy change. Errors for ΔC_p were calculated using the Microsoft Excel Analysis ToolPak linear regression. The isentropic temperatures were determined from linear fits of the entropic contribution (−TΔS). All presented thermodynamic parameters were derived from these ITC series. Parameter errors were determined using a Monte Carlo approach using the errors derived from the temperature dependence linear fits. Reported ITC stoichiometry (N-values) represents mean and standard deviations of all experiments.

**NMR Spectroscopy.** NMR data were acquired on Bruker AVANCE 750 or 800 MHz (1H) spectrometers equipped with cryogenic probes. Free induction decays were transformed and visualized using NMRPipe and analyzed using the Gn/Nmr Analysis software. All samples were prepared using 200–300 μM 15N- or 13C,15N-labeled protein with, if relevant, a 100 μM excess of the unlabeled binding partner. All samples were prepared using 10% (v/v) D_2O, 0.02% (w/v) NaNO_3, 0.2 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), 20 mM NaHPO_4/NaH_2PO_4 (pH 7.0), and 100
mM NaCl. Backbone assignments of DREB2A_{234-276} in the free and bound state were performed manually from analysis of $^{1}H-^{13}C-^{15}N$-HSQC, HNCA, HNCO, HNCOACB, and HNCAHNCO spectra. Assignments of backbone resonances of RCD1-RST in complex with DREB2A_{234-276} were done by analysis of $^{1}H-^{13}C-^{15}N$-HSQC, HNCA, and HNCOACB spectra. Resonances for RCD1-RST in the free form and in complex with DREB2A_{234-276} were acquired from previous work.\textsuperscript{19} CSP analysis was done with combined backbone amide group chemical shift changes using a 0.154 weight for the $^{15}N$ shifts.\textsuperscript{68} HNCANNH spectra. Assignment so fb a c k b o n e r e s o n a n c e so f HSQC, HNCACB, HNCOCACB, HNCO, HNCOAC, and HNCAHNCO spectra. Chemical shifts and $^{15}N$ relaxation rates for the unbound DREB2A_{234-276} were done by using phi-psi angles predicted from NMR chemical shifts using TALOS-N.\textsuperscript{18} The complex structure models were generated using CABS-dock\textsuperscript{71} and FlexPepDock\textsuperscript{72} using intermolecular contacts derived from published structures of the RCD1-RST:DREB2A_{234-276} complex and secondary structure predicted from NMR chemical shifts.

## ASSOCIATED CONTENT

\section*{Supporting Information}

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.1c04214.

Tables S1–S4, Figures S1–S8, Script 1 (PDF)

## AUTHOR INFORMATION

\section*{Corresponding Authors}

Birthe B. Kragelund \textsuperscript{#} – Biomolecular Sciences and Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, 2200 Copenhagen N, Denmark; \url{orcid.org/0000-0002-7454-1761}; Email: bbb@bio.ku.dk

Karen Skriver \textsuperscript{#} – Biomolecular Sciences, University of Copenhagen, 2200 Copenhagen N, Denmark; \url{orcid.org/0000-0003-2225-4012}; Email: kskriver@bio.ku.dk

\section*{Authors}

Frederik Friis Theisen \textsuperscript{#} – Biomolecular Sciences and Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, 2200 Copenhagen N, Denmark; \url{orcid.org/0000-0002-3412-4242}

Lasse Staby \textsuperscript{#} – Biomolecular Sciences and Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, 2200 Copenhagen N, Denmark; \url{orcid.org/0000-0002-5588-5163}

Frederik Gronbæk Tidemand \textsuperscript{#} – Structural Biophysics, X-ray and Neutron Science, The Niels Bohr Institute, University of Copenhagen, 2100 Copenhagen \O, Denmark; \url{orcid.org/0000-0001-8914-9626}

Charlotte O’Shea \textsuperscript{#} – Biomolecular Sciences, University of Copenhagen, 2200 Copenhagen N, Denmark; \url{orcid.org/0000-0002-5548-8002}

Andreas Prestel \textsuperscript{#} – Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, 2200 Copenhagen N, Denmark; \url{orcid.org/0000-0003-1689-2712}

\section*{Complete contact information is available at:}

https://pubs.acs.org/10.1021/jacs.1c04214

\section*{Funding}

This work was funded by the Novo Nordisk Foundation Challenge Grant REPIN (\#NNF18OC0033926 to B.B.K. and K.S.) and Novo Nordisk Foundation grant (\#NNF18OC0052177 to K.S.). NMR data were recorded at cOpenNMR, an infrastructure facility funded by the Novo Nordisk Foundation (\#NNF18OC0032996).

\section*{Notes}

The authors declare no competing financial interest.

Chemical shifts and $^{15}N$ relaxation rates for the unbound DREB2A_{234-276} (accession code S1055) and for the RCD1-RST:DREB2A_{234-276} complex (accession code S1056) have been deposited in the BMRB.

\section*{ACKNOWLEDGMENTS}

Prof. Krenst Lindorff-Larsen and Drs. Katrine Bugge and Yulian Gavrilov are thanked for useful discussions.

\section*{REFERENCES}


