A phosphorylation-motif for tuneable helix stabilisation in intrinsically disordered proteins - Lessons from the sodium proton exchanger 1 (NHE1)

Hendus-Altenburger, Ruth; Lambrughi, Matteo; Terkelsen, Thilde Bagger; Pedersen, Stine Helene Falsig; Papaleo, Elena; Lindorff-Larsen, Kresten; Kragelund, Birthe Brandt

Published in:
Cellular Signalling

DOI:
10.1016/j.cellsig.2017.05.015

Publication date:
2017

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

Cellular Signalling 37 (2017) 40–51

Contents lists available at ScienceDirect

Cellular Signalling

journal homepage: www.elsevier.com/locate/cellsig

A phosphorylation-motif for tuneable helix stabilisation in intrinsically disordered proteins – Lessons from the sodium proton exchanger 1 (NHE1)

Ruth Hendus-Altenburger[a], Matteo Lambrughi[b], Thilde Terkelsen[b], Stine F. Pedersen[c], Elena Papaleo[b], Kresten Lindorff-Larsen[a,*], Birthe B. Kragelund[a,*]

[a] Structural Biology and NMR Laboratory, Linderstrøm-Lang Centre for Protein Science, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark
[b] Computational Biology Laboratory, Center for Autophagy, Recycling and Disease, Strandboulevarden 49, 2100 Copenhagen, Denmark
[c] Cell Biology and Physiology, Department of Biology, University of Copenhagen, Universitetsparken 13, DK-2100 Copenhagen Ø, Denmark

ARTICLE INFO

Keywords:
IDP
SLiM
Phospho-cap
Secondary chemical shifts
NMR

ABSTRACT

Intrinsically disordered proteins (IDPs) are involved in many pivotal cellular processes including phosphorylation and signalling. The structural and functional effects of phosphorylation of IDPs remain poorly understood and difficult to predict. Thus, a need exists to identify motifs that confer phosphorylation-dependent perturbation of the local preferences for forming e.g. helical structures as well as motifs that do not. The disordered distal tail of the Na^+/H^+ exchanger 1 (NHE1) is six-times phosphorylated (S693, S723, S726, S771, T779, S785) by the mitogen activated protein kinase 2 (MAPK1, ERK2). Using NMR spectroscopy, we found that two out of those six phosphorylation sites had a stabilizing effect on transient helices. One of these was further investigated by circular dichroism and NMR spectroscopy as well as by molecular dynamic simulations, which confirmed the stabilizing effect and resulted in the identification of a short linear motif for helix stabilisation: [S/T]-P-{3}-[R/K] where [S/T] is the phosphorylation-site. By analysing IDP and phosphorylation site databases we found that the motif is significantly enriched around known phosphorylation sites, supporting a potential wider-spread role in phosphorylation-mediated regulation of intrinsically disordered proteins. The identification of such motifs is important for understanding the molecular mechanism of cellular signalling, and is crucial for the development of predictors for the structural effect of phosphorylation; a tool of relevance for understanding disease-promoting mutations that for example interfere with signalling for instance through constitutive active and often cancer-promoting signalling.

1. Introduction

Phosphorylation is one of the main mechanisms underlying cellular signalling. The transduction of phosphorylation-mediated signals is rooted in the modulation of molecular interactions, which results from changes to the chemical structure. A chemical modification can lead to more than just chemical change, for example by inducing local structural effects or even global ones. Thus, phosphorylation sites that are important for modulating the interaction network can be divided into those that only require a charge change for recognition, and those that additionally modulate the local/global structure (structural motifs). Currently, however, it is not possible to differentiate between these cases and the molecular details underlying the events are not clear.

1.1. Intrinsically disordered proteins and phosphorylation

Intrinsically disordered proteins (IDPs) often transition between disordered and ordered states to mediate their biological function via binding to partner proteins. In many cases binding is accompanied by local folding, either by conformer selection or induced fit, or a mixture of the two [1–5]. The propensity to form helical structures can accelerate ligand binding [6], and phosphorylation can regulate local and/or global folding, and thus biological function, shown e.g. for the disordered eukaryotic translation initiation factor 4E-binding protein 2, where multisite phosphorylation leads to folding [7]. In yeast Ash1, however, compensatory structural effects make overall chain expansion insensitive to phosphorylation [8]. The effect of phosphorylation on local structure depends on several factors, i.e. presence and type of

Abbreviations: NHE1, Na^+/H^+ exchanger 1; NHE1cdt, C-terminal distal tail of NHE1; CS, chemical shift; SCS, secondary chemical shift; PTM, posttranslational modification

* Corresponding authors.
E-mail addresses: rha@bio.ku.dk (R. Hendus-Altenburger), matl@cancer.dk (M. Lambrughi), thilde@cancer.dk (T. Terkelsen), sfpedersen@bio.ku.dk (S.F. Pedersen), elenap@cancer.dk (E. Papaleo), lindorff@bio.ku.dk (K. Lindorff-Larsen), bbk@bio.ku.dk (B.B. Kragelund).

http://dx.doi.org/10.1016/j.cellsig.2017.05.015
Received 31 January 2017; Received in revised form 24 May 2017; Accepted 25 May 2017
Available online 26 May 2017

transient secondary structure elements, the relative position of the phosphorylation site to preformed structural elements [9,10], the neighbouring sequence such as a hydrogen bond network and electrostatic interactions [11], the local environment of the protein including the presence of aromatic residues [12,13], as well as the number of phosphorylation sites [14]. Furthermore, the inductive effect is dependent on the type of phosphorylated residue, i.e. Ser, Thr, or Tyr [15]. Thus, the concerted contributions of all those factors determine the net modulation in local structure, but as of yet, the origins and relative importance of these effects are not fully understood and remain difficult to predict.

1.2. Structural preferences from NMR chemical shifts

The most direct NMR method to identify secondary structures in proteins is based on the secondary chemical shift (SCS) analysis (Δδ). Chemical shifts (δ, CSs) of the backbone nuclei and in particular those of the H^{α}, C^{α} and C' correlate strongly with local backbone structure and by comparison to random coil CSs one can derive secondary structure and specifically identify the position, length and population of those in proteins. SCSs are defined as $\Delta\delta = \delta(\text{observed}) - \delta(\text{random coil})$. Specifically, consecutive positive secondary C^{α} and C' CSs indicate helicity, negative values β-sheet or extended structure and close to zero values indicate that the ensemble populates all allowed regions in the Ramachandran plot and hence appear random [16,17]. Yet, random coil CSs of individual residues vary depending on neighbouring residues [18] as well as the experimental conditions (pH, temperature) [19]. Currently, several datasets for reference CSs of naturally occurring amino acids exist that use either neighbouring correction factors derived from peptide libraries [19–21] or a curated chemical shift data set derived from disordered proteins [22]. Secondary shift analysis is commonly used to assess transient structure in IDPs. Yet, the transient nature of secondary structures in IDPs results in small secondary chemical shift values and makes structure analysis more prone to bias by the exact reference values used for the random coil shifts. In addition, experimental conditions as well as posttranslational modifications (PTMs) affect the CSs, and the chemical effects of PTMs, like phosphorylation, on the random coil CSs are not yet included in these libraries. Previous studies have aimed to characterize the random coil CSs of phosphorylated amino acids using peptides [23]. However, no correction factors for modification of neighbouring amino acid residues are currently available and the effect of a PTM may extend further than just to its immediate neighbours. Additionally, in case of secondary structure modulation by amino acid modification, it is currently not possible to separate the chemical effects from structural effects. One approach is to use intrinsic random coil CSs assigned in 8 M urea to determine the SCSs ($\Delta\delta_{irc}$) [24]. The exact mechanism by which urea unfolds proteins is not fully understood, but it is assumed that urea promotes backbone disorder [25] and urea has been shown to promote poly-proline II conformations [26,27]. Yet, the use of urea allows the study of the unfolded state at physiological pH and specific ionic strength. The use of the chemically unfolded state as internal reference does not require a reference library, as each protein can be used as its own internal reference and the approach has successfully been used for characterisation of the acid unfolded state [24,28,29] as well as for secondary structure analysis for a number of IDPs [30,31].

1.3. The Na^+/H^+ exchanger isoform 1 and its disordered tail

The Na^+/H^+ exchanger isoform 1 (NHE1) plays a central role in the regulation of cellular pH and volume, and in turn NHE1 is important for the normal physiological functions of most mammalian tissues [32,33]. Its function is tightly regulated through mechanisms involving interactions with multiple protein- and lipid-binding partners, and through phosphorylations, and other PTMs. The intracellular distal tail of NHE1 (NHE1cdt) is intrinsically disordered [34] and contains most of the many known and suggested phosphorylation sites [32]. In vitro, NHE1cdt is six times phosphorylated (S693, S723, S726, S771, T779, and S785) by the mitogen activated protein kinase 1 (MAPK1, ERK2), and although the in vivo role of ERK2 in direct phosphorylation of NHE1 is contentious, NHE1 engages in regulation of ERK2 activity and NHE1 is regulated by ERK signalling [31,35,36]. Thus, NHE1cdt is an excellent system for investigation of the potential structural effects of Ser/Thr-phosphorylation in IDPs.

Here, using NHE1 as a model protein, we asked what the structural details underlying phosphorylation-induced structural effects in IDPs would be and used the intrinsic random coil shifts to characterize the secondary structure profile of NHE1cdt and the modulation of its transient structure by the six phosphorylations. Phosphorylation at four of the six sites did not induce any significant change to the structure profile, yet two of the phosphorylations led to a substantially increased local propensity to form helical structure. For one of these sites, S785, CD and NMR spectroscopy and molecular dynamic simulations further confirmed this effect. These results allowed us to discover a specific motif that we suggest confers phosphorylation-induced N-terminal helix stabilisation that we further found is generally enriched in IDPs and at confirmed phosphorylation sites. This motif and its corresponding tuneable behaviour induced by phosphorylation will be important for understanding how the interactome of IDPs can be modulated. Furthermore, the identification of stabilizing motifs is essential to develop bioinformatics predictors for the effect of phosphorylation, which may be of value for understanding disease related mutations in general.

2. Materials and methods

2.1. Expression, purification and phosphorylation of hNHE1cdt

Expression, purification and phosphorylation of unlabelled, ^{15}N-labelled and ^{13}C,^{15}N-double labelled hNHE1cdt, as well as its phosphorylation by active ERK2 was performed essentially as described [31]. The phosphorylation reaction was stopped and the kinase removed by heating the sample to 90 °C for 5 min followed by 20,000*g* centrifugation for 5 min. Phosphorylation of all six sites was confirmed by mass spectrometry and by analyses of a ^{15}N,^{1}H-HSQC NMR spectrum.

2.2. NMR samples and spectral acquisition

All NMR spectra were recorded on either a Varian INOVA 750 MHz, a 800 MHz ^{1}H NMR spectrometer equipped with a 5 mm triple resonance probe with a Z-field gradient, or a 600 MHz Bruker Avance III HD spectrometer in 5 mm Shigemi NMR tubes. Backbone resonance assignments of hNHE1cdt in the unphosphorylated and phosphorylated states were done at 5 °C using 0.5–1.5 mM samples of ^{15}N,^{13}C-hNHE1cdt in PBS pH 7.2 supplemented with 0.5 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), 10 mM dithiothreitol (DTT), and 10% (v/v) 99.96% D_2O by standard 3D triple resonance experiments as previously described [34]. Intrinsic random coil referencing was done from assignments of hNHE1cdt in 8 M urea from similar experiments. The 8 M urea samples were prepared as follows: the protein was dialysed against 2 × concentrated PBS buffer, and then 200 µL protein stock mixed with 192.3 mg urea (spectral grade), supplemented with 0.5 mM DSS, 10 mM DTT and 10% (v/v) 99.96% D_2O as above, pH adjusted, and water added to a final volume of 400 µL. Chemical shift referencing was done relative to DSS. The spectra were zero filled, apodized, Fourier transformed and baseline corrected in NMRDraw [37] and analysed manually using CCPN Analysis [38]. The secondary structure analysis was done by secondary chemical shifts, calculated as

$$\Delta\delta_{irc} = \delta - \delta_{irc} \quad (1)$$

or as

$$\Delta\delta_{rc} = \delta - \delta_{rc} \quad (2)$$

using two different random coil shift sets [19,22], and the SSP score [20], as well as Agadir predictions [39].

The changes in $^{1}H^{N}$ chemical shifts with temperature were obtained from recording a series of $^{15}N,^{1}H$-HSQCs from 4-° to 16 °C with increments of 2 °C. Temperature coefficients and standard deviations were extracted with Matlab by fitting the $^{1}H^{N}$ chemical shifts against temperature using the following linear equation:

$$y = a*x + y_0 \quad (3)$$

where y is the $^{1}H^{N}$ chemical shift (ppb), a is the temperature coefficients (ppb/°C), x is the temperature (°C), and y_0 is the offset. The NHE1 TH4 peptides TH4pepWT, TH4pepD784A, and TH4pepR790V with and without S785 phosphorylation were purchased from Synpeptide Co. Ltd. (China) and Schafer-N (DK) purified by rpHPLC to > 95% purity and confirmed by mass spectrometry. The NMR resonances of the protons, N^{H} and C^{α} were assigned from homonuclear DQF-COSY, TOCSY, ROESY spectra, as well as natural abundance $^{15}N,^{1}H$- and $^{13}C,^{1}H$-HSQCs at 5 °C. The salt titration was done by stepwise addition of NaCl to the peptide samples and subsequent acquisition of $^{13}C,^{1}H$-HSQCs. Percent helicity was estimated from the assigned secondary chemical shifts using the intrinsic random coil shifts recorded in urea, averaged over the length of the transient helices and divided by 2.8 ppm, the secondary C^{α} chemical shift expected for a 100% formed helix [40].

2.3. Molecular dynamics simulations

We extracted the starting structure to use in molecular simulations from an ensemble of 1000 disordered conformations of the unphosphorylated NHE1 TH4pep construct generated with *Profasi* [41] at 300 K and the phosphorylated variant was modelled using the FoldX-suite without any repair of the backbone conformation upon the substitution [http://foldxsuite.crg.eu/]. To generate the ensemble of disordered conformations, we employed a standard Monte Carlo move set, as used in earlier studies with *Profasi* [42]. In particular, we selected a structure from the ensemble with the peptide in an extended conformation to avoid artificial tertiary contacts. We performed the simulations with Gromacs 4.6 [43]. The system was solvated in a dodecahedral box of water molecules at 150 mM NaCl, applying periodic boundary conditions with a minimum distance of 15 Å between the peptide and the box edges. We used the Amber03ws protein force field [44] recently developed for disordered proteins.

We used parameters specifically developed for the amber force field [45] to describe the phosphorylated serine. Simulations were carried out at 298 K using the canonical NVT ensemble. We used the LINCS algorithm to constrain heavy-atom bonds, allowing for a 2 fs time-step. Long-range electrostatic interactions were described by the Particle-mesh Ewald summation scheme. Van der Waals and Coulomb interactions were truncated at 9 Å. After a series of energy minimization, solvent equilibration, thermalization and pressurization preparatory steps, we performed a one-microsecond long MD simulation. To define a hydrogen bond (H-bond) we used empirical geometric criteria. In particular, we defined the presence of an H-bond if the distance between the donor and acceptor atoms is less or equal than 3.0 Å (relaxed cutoff) or 2.5 Å (strict cutoff) and the angle between the donor atom-hydrogen-acceptor atom is equal or higher than 90° [46]. The helical content in the MD simulations was calculated according to DSSP definition [47]. Specifically, we considered as helical residues collectively those that DSSP assigned either as 3_{10} helix, α-helix or π-helix.

2.4. Motif enrichment analysis

2.4.1. Query set and background set of disordered regions for motif enrichment

Datasets of intrinsically disordered proteins were obtained from the Database of Disordered Protein Predictions [48]. The following files were downloaded from $D^{2}P^{2}$ on the 6th of January 2016: d2p2_protein_to_uniprot.tsv, consensus_ranges.tsv, dis_experiment.tsv and ptm_assignment.tsv. Consensus regions of disorder obtained from prediction tools (consensus_ranges.tsv) were mapped to their UniProt identifier using the file d2p2_protein_to_uniprot.tsv. Regions labelled as disordered by experiments were extracted from the file dis_experiment.tsv. If UniProt IDs were found in both predicted and experimental set, regions from the predicted set were removed in order to avoid redundancy and to keep only the experimental entry. Consensus regions from prediction and experimental disordered regions were merged after which all regions shorter than 10 amino acid residues were removed from the dataset. In order to obtain only disordered regions from eukaryotic proteins, a list of species (speclist) with assigned taxon identifier and domain was obtained from uniprot (http://www.uniprot.org/docs/speclist, downloaded on 12th of January 2016). Taxon IDs from the speclist were then used to obtain the UniProt accession numbers of eukaryotic proteins from the NCBI resource prot.accession2taxid (ftp://ftp.ncbi.nih.gov/pub/taxonomy/accession2taxid/, downloaded on 12th of January 2016).

All disordered regions (experimental + predicted) not assigned to any eukaryotic UniProt accession number were removed from the analysis. Known phosphorylated serines and threonines were extracted from ptm_assignment.tsv and were subsequently merged with the disorders regions (experimental + predicted) using the internal $D^{2}P^{2}$ IDs. Disordered regions encompassing a phosphorylated serine or threonine were extracted resulting in a query dataset for motif enrichment. The query dataset consisted of a total of 21,653 disordered regions (10,268 unique UniProt IDs). The background set for enrichment analysis included all disordered regions containing at least one serine or threonine without any known post-translational phosphorylation resulting in a total of 155,416 disordered regions (54,420 unique UniProt IDs). FASTA-format sequences of both query - and background regions were obtained using a custom R-script with the R-packages protr [49] and seqRFLP [50].

2.4.2. Motif enrichment with AME (analysis of motif enrichment)

Enrichment analysis was performed using AME (Analysis of Motif Enrichment) [51], which is a part of the MEME suit of motif discovery tools [52]. The two sets of sequences, those with phospho-serine/threonine and those without (but containing at least one serine/threonine) were merged in order to calculate the background frequencies of amino acid residues. This calculation was performed using the MEME tool fasta-get-markov (https://github.com/CPFL/gmeme/blob/master/meme_4.8.1/doc/fasta-get-markov.xml). The calculated residue frequencies were used to construct a MEME motif file containing custom motif probability matrices. The frequencies of the three "variable" residues within the query motif (denoted X) were set to either to uniform frequencies of 0.05, or to those of the calculated background frequencies (See weighed matrixes in Supplementary information File S8). In addition to the motif of interest enrichment was performed using the four known (non-structural) phospho-motifs of NHE1cdt with similar length (approximately seven amino acid residues) A-S-P-{3}-E, S-S-P-{3}-D, Q-S-P-{3}-D and F-T-P-{3}-D (See weighed matrixes in Supporting information File S8). Enrichment with AME was performed using Fisher's exact test and enrichment included correction for sequence length [51]. Cut-off for p-value to be considered significant was ≤ 0.01.

Table 1
C^{α}, C^{β} and C′ CSs of the six phosphorylation sites of NHE1cdt with and without phosphorylation.

	C^{α} (ppm)			C^{β} (ppm)			C′ (ppm)		
	WT	Phos	Δ	WT	Phos	Δ	WT	Phos	Δ
S693	56.44	55.92	− 0.52	63.28	65.15	+ 1.87	173.08	173.72	+ 0.64
S723	56.47	55.89	− 0.58	63.15	64.55	+ 1.40	172.72	172.34	− 0.38
S726	56.40	56.16	− 0.24	63.13	64.54	+ 1.41	172.92	172.27	− 0.65
S771	56.46	56.00	− 0.46	63.21	64.73	+ 1.52	172.65	172.70	+ 0.05
T779	59.31	60.34	+ 1.03	69.96	72.72	+ 2.76	171.81	172.06	+ 0.25
S785	56.50	55.79	− 0.71	63.26	65.06	+ 1.80	173.03	173.66	+ 0.63

2.4.3. Motif enrichment with chi-squared test in R

In addition to AME, a simpler approach to enrichment was implemented in R. Enrichment analysis was performed using a regular expression to capture all motifs in both query and background sets. The motif counts were used to perform a χ^2 test of significance and to calculate the odd ratio (OR). Cut-off for significance was a false discovery rate of 0.01. After enrichment, the location of the serine/threonine residue in the motifs in the query set was extracted and matched with the position of the PTM serines/threonines.

2.5. Circular dichroism spectroscopy

Far-UV CD measurements were recorded on a Jasco J-810 spectropolarimeter with Peltier control at 10 °C using a 1 mm path length, from 260 to 190 nm with a scan speed of 50 nm/min, 10 accumulations, and with a data pitch of 0.1 nm. Peptide concentrations of all TH4pep variants were 70 μM. Sample conditions were 20 mM sodium phosphate buffer, 1 mM DTT, pH 7.2. Spectra were buffer corrected and converted to mean residual ellipticity using the equation:

$$[\theta]_{MR} = 100*\theta/(c*l*n) \quad (4)$$

where θ is the ellipticity (deg), c is the peptide concentration (M), l is the path length of the cuvette (cm), and n is the number of amino acid residues.

2.6. MALDI-TOF mass spectrometry

One microliter of 10 μM NHE1cdt before and after phosphorylation was mixed with 1 μL HCCA matrix [10 mg/mL in 50% (v/v) MeCN and 0.1% (v/v) TFA] and spotted on a MTP 384 ground steel TF target plate by the dried droplet method. Spectra were recorded on a Bruker Daltonics Autoflex II TOF mass spectrometer in linear positive mode. An average of 100 shot was recorded. Calibration was performed with quadric calibration of protein calibration standard 1 (Bruker Daltonics).

3. Results

3.1. The transient secondary structure profile of the hNHE1cdt

Secondary structure analysis by NMR spectroscopy revealed several transient helices (THs) in the intrinsically disordered NHE1cdt, initially identified using the SSP score [34] and refined by the SCSs using the intrinsic random coil reference CSs obtained in 8 M Urea ($\Delta\delta_{irc}$) [31]. These were located at P694-R700 (TH1), P727-G743 (TH2), D758-S766 (TH3), and P786-L795 (TH4), where TH1 and TH4 were rather short and lowly populated (Fig. 1A–B). Using a SCS value of 2.8 ppm for C^{α} to represent that of a 100% populated helix [40], the transient helices were estimated to be populated to 7.3% (TH1), 11% (TH2), 7.9% (TH3) and 5.6% (TH4). A comparison of the $\Delta\delta_{irc}$ with $\Delta\delta_{rc}$ derived from different random coil libraries and datasets as well as Agadir predictions is shown in Supplementary Fig. S1. The $\Delta\delta_{rc}$ derived from different random coil datasets predicted the THs in the NHE1cdt with comparable intensities and positions. However, the noise level was considerably higher compared to that obtained using the $\Delta\delta_{irc}$ (Supplementary Fig. S1B–C). The SSP score and Agadir also predicted similar positions, but the SSP score had negative SCS regions not present in the $\Delta\delta_{irc}$, and Agadir generally underestimated the amplitude of the helical propensity (Supplementary Fig. S1D–E). Since the random coil shifts of phosphorylated libraries are unavailable and as the intrinsic reference coil method is able to extract the pure chemical effect of the phosphorylation on the chemical shift, we proceeded with this method.

3.2. Modulation of local transient structure of the hNHE1cdt by phosphorylation

The NHE1cdt contains six ERK2 consensus phosphorylation sites that are phosphorylated by active ERK2 in vitro (S693, S723, S726, S771, T779 and S785) [31]. All six ERK2 sites in hNHE1cdt are located in “linker” regions or directly N-terminal to the transient helices. Specifically, S693 is located N-terminal to TH1, S723 and S726 N-terminal to TH2, S771 and T779 locate to the linker region between TH3 and TH4, and S785 is located N-terminally to TH4 (Fig. 1A–B). To determine the structural effects of these phosphorylations, sets of backbone spectra were recorded for the six-times phosphorylated state under the same experimental conditions as for the unphosphorylated state with and without 8 M urea to obtain intrinsic reference coil shifts. The $\Delta\delta_{irc}$ for the phosphorylated state showed similar profile as for the unphosphorylated state, but with increased helicity for TH1 and TH4 (Fig. 1C). Since phosphorylation can have a structural effect but also affect the random coil shift due to the chemical modification, the $\Delta\delta_{irc}$ will automatically correct for the chemical effect, as this is expected to be the same in the presence and absence of urea. To illustrate this further the chemical shifts of the unphosphorylated and phosphorylated state were compared in Fig. 1D (structural and chemical effect), as well as both of them in urea (chemical effect only, Fig. 1E). Comparing the C^{α} CSs for the phosphorylated with those of the unphosphorylated state revealed a downfield shift for all phospho-serines, and an upfield shift for T779 (Table 1, Fig. 1D). Also, many of the immediate C-terminal neighbours experienced a C^{α} downfield shift and N-terminal neighbours an upfield shift. The C′ shifts revealed a less systematic pattern, with mixed up- and downfield shifts for the phosphorylated residues (Table 1) and for the immediate neighbours (Fig. 1D). To extract the structural effect of phosphorylation, the $\Delta\delta_{irc}$ of the unphosphorylated state was subtracted from the $\Delta\delta_{irc}$ of the phosphorylation state ($\Delta\Delta\delta_{irc}$, Fig. 1F), showing a small increase of helicity of TH1 by S693 phosphorylation (from 7.3% to 13%) as well as a stronger increase in TH4 by S785 phosphorylation (increasing from 5.6% to 16%). Phosphorylation at the other sites did not appear to cause any substantial structural effects. The separation of structural from chemical effect for TH4 is shown in Fig. 1G, which highlights that the shifts of S783-S785 were mainly due to the chemical modification, and that the shifts of P786-S788 had an almost equal contribution from chemical and structural effects. Structural effect thus dominated the shifts further away from the phosphorylation sites.

S693 and S785 are located N-terminally to TH1 and TH4, respectively, and might thus stabilize these via interactions with a helix

dipole. However, S723 and S726 are also located just N-terminally to TH2, which is not stabilized by phosphorylation (from 11% to 10% in the two states), demonstrating that phosphorylation just N-terminally to a helix is not always sufficient to increase its stability. These observations thus suggest that either other mechanisms contribute to the observed stabilisation of TH1 and TH4 or possibly that a helix-dipole based mechanism requires a substantial population of helix prior to phosphorylation. In order to find alternative potential mechanisms we scrutinized the sequences and structures around the phosphorylation sites and discovered that the sequences surrounding S693 and S785, whose phosphorylation stabilize TH1 and TH4, respectively, share common sequence properties (Fig. 1H). In particular, relative to the phosphorylation site, *i*, both sites have an Asp at the $i - 1$ and an Arg at the $i + 5$ position, features that are not present at any of the other four sites, and we thus explored whether these might contribute to the structural response to phosphorylation.

3.3. Design of a NHE1 peptide for NMR and molecular dynamics investigation

To provide a more detailed view, and to evaluate the potential role of the Asp and Arg residues for the stabilisation, we next performed further NMR experiments and molecular simulations on a peptide covering the S785 phosphorylation site and the transient helix TH4 (TH4pep, $NHE1_{781–796}$). This peptide, rather than that surrounding S693, was used for several reasons: (i) the S693 phosphorylation site has a His in close proximity (H698) and given the pH-sensitivity of histidines, even minor changes in pH would affect the result, (ii) S693 has a Lys in the i − 3 position and a basic residue in this position was previously shown to stabilize helicity [53–56], and (iii) the stabilizing effect was stronger for the S785 phosphorylation site (Fig. 1F).

3.4. Decomposing the individual effects using spectroscopy

To evaluate the potential individual roles of the Asp and Arg for the stabilisation we used peptide variants. First, using Agadir we screened for mutations at the two positions, D784 and R790 that according to Agadir would retain the helicity of the peptide but replace the charges (Supplementary Fig. S2). This resulted in two peptide-based variants: TH4pepD784A and TH4pepR790V. We next used circular dichroism (CD) spectroscopy to evaluate the effect of those mutations on the phosphorylated versus unphosphorylated states of the peptides (Fig. 2A). While for all three variants, including the WT peptide, the ellipticity at 190 nm increased upon phosphorylation, indicating less disorder, the negative ellipticity at 222 nm, indicating an increase in helicity, was only increased for the WT and the D784A variants. Thus, although we cannot at this point explain the decrease at 190 nm, the change at 222 nm clearly indicated R790 to be important for conferring the phosphorylation-induced helix stabilisation. Furthermore, the apparent discrepancy between the absolute level of helicity measured by NMR and by CD is not surprising as fragmentations of helices such as dynamically formed transient helices lowers the ellipticity from the *n*-π* transition in CD spectroscopy [57].

To assess whether only the local sequence is responsible for the stabilizing effect of phosphorylation and whether it is independent of the nearby T779 phosphorylation we next assigned the NMR chemical shifts of the TH4pepWT with and without phosphorylation by natural abundance NMR spectroscopy. In the $^{1}H,^{15}N$-HSQC spectrum, the amide peaks for the TH4pep have similar shifts as those within NHE1cdt, indicating that TH4 behaves similar in the peptide as in the NHE1cdt (Fig. 2B and Supplementary Fig. S3). Furthermore, in the $^{1}H,^{15}N$-HSQC of the TH4pep the Arg side chain peaks were visible and folded into the spectra (Fig. 2B). The $H^{\eta}N^{\eta}$ peaks of R790 and R793 were almost unaffected by phosphorylation. However, the $H^{\varepsilon}N^{\varepsilon}$ peak of R790 was clearly and significantly affected by phosphorylation. While the $H^{\varepsilon}N^{\varepsilon}$ peaks of R790 and R793 overlapped in the unphosphorylated state, the $H^{\varepsilon}N^{\varepsilon}$ peak of R790 moved away from the $H^{\varepsilon}N^{\varepsilon}$ peak of R793 in the phosphorylated state (Fig. 2B **zoom**), further confirming its role in the stabilisation.

We noticed that the effect of phosphorylation on the C^{α} chemical shifts of TH4 was slightly higher for the TH4pep compared to TH4 in NHE1cdt (Fig. 2C) and speculated whether this was due to the absence of salt in the peptide sample, which suggests the nature of the stabilizing effect to be at least in part electrostatic. To substantiate this we recorded $^{1}H,^{13}C$-HSQC spectra of both unphosphorylated and phosphorylated TH4pep at increasing salt concentrations and followed the C^{α} chemical shifts (Fig. 2D–E). While the chemical shifts of the unphosphorylated TH4pep were essentially unaffected (as exemplified by P786, S788, and R790 in Fig. 2E), the chemical shifts of the phosphorylated peptide were substantially affected by salt (Fig. 2E). Thus, at high salt concentration the phosphorylation-induced helix stabilisation was much weaker than at low ionic strength (Fig. 2D). This suggests strongly that the stabilizing interactions have a large electrostatic component.

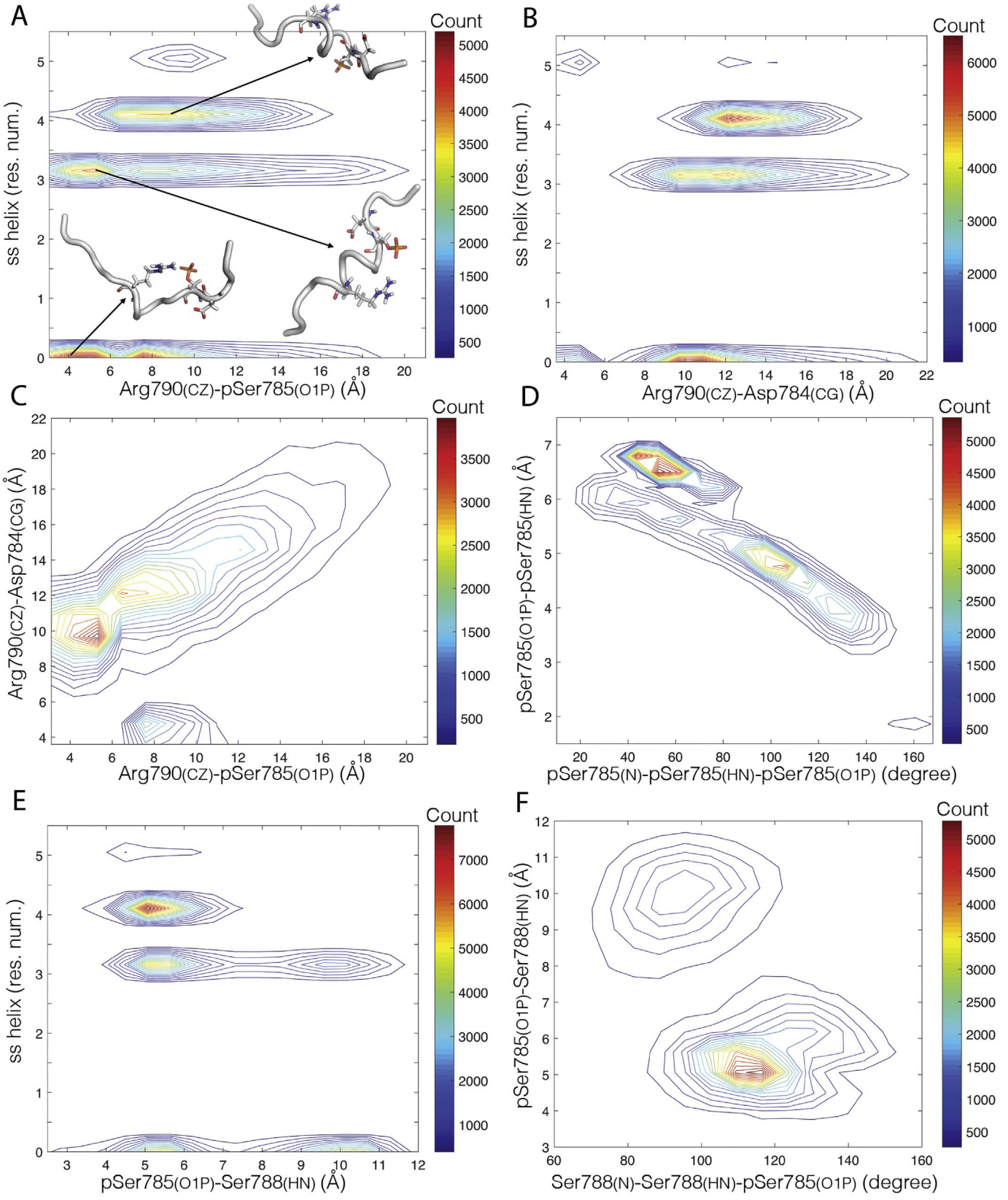
In conclusion, these data show that R790 is important for the stabilizing effect upon S785 phosphorylation, and involves the $H^{\varepsilon}N^{\varepsilon}$ in the Arg side chain that could putatively form of a hydrogen bond to the phosphoryl-group. Although electrostatics play a significant role in the stabilisation the Asp appears not to play any role.

3.5. Temperature coefficients and hydrogen bond donors

The strong upfield shift in the $^{15}N,^{1}H$-HSQC for phosphorylated amino acid residues has been suggested to be due to hydrogen bond formation between the phosphorylated side chain and its own backbone amide [58,59]. However, in 8 M urea, we still observed an upfield shift for the phosphorylated residues suggesting either that no hydrogen bond was initially formed or that it is preserved in urea (Supplementary Fig. S3D–E). Conveniently, hydrogen bond donors can be detected by measuring temperature coefficients for the amide protons, where the chemical shifts of hydrogen-bonded amides generally are less dependent on temperature. In particular, amides with temperature coefficients more positive that − 4.6 ppb/°C are generally hydrogen bonded (85% of hydrogen-bonded residues fall in this range), with the predictive value increasing to ~93% when the temperature coefficient is greater than − 4.0 ppb/°C [60].

To identify hydrogen bonding in NHE1cdt we recorded a series of $^{15}N,^{1}H$-HSQCs for both the unphosphorylated and phosphorylated state with and without 8 M urea as a function of temperature. The temperature coefficients for TH4 are shown in Fig. 2F–G, whereas the coefficients for the entire NHE1cdt can be found in Supplementary Fig. S4. The only amide of unphosphorylated TH4pep that is clearly above the threshold of − 4.0 ppb/°C is S796 (~ − 3 ppb/°C), whereas Q789 and R790 fall in the range between − 4.6 ppb/°C and − 4.0 ppb/°C, suggesting that they potentially engage in hydrogen bonds with close by residues stabilizing the transient helix. Upon phosphorylation, all of the amide temperature coefficients were essentially unaffected, except for the very large increase for S788 (− 5.2 ppb/°C to − 1.7 ppb/°C), suggesting a switch in hydrogen-bonding status upon phosphorylation (Fig. 2F). S788 also showed a strong upfield shift in the $^{15}N,^{1}H$-HSQC upon phosphorylation (Fig. 2H). In urea the temperature coefficients were much lower confirming the absence of hydrogen bonds in the reference state (Fig. 2G).

Collectively, these data indicated that phosphorylation might lead to hydrogen bond formation involving an amide of the backbone. Interestingly, the amide of the pS785 was not affected, i.e. hydrogen bond formation was not observed before or after phosphorylation. This observation suggests that the strong downfield shift of an amide in an $^{15}N,^{1}H$-HSQC of the phosphorylated residue itself is not due to its engagement in a hydrogen bond. Rather the downfield shift appeared to be simply due to the additional charge(s) added by the phosphate group (− 2 at physiological pH), i.e. de-shielding due to the presence of a



(caption on next page)

Fig. 3. Phosphorylation of S785 has an impact on the conformational ensemble of TH4pep. The analyses were performed on an unbiased all-atom MD simulation of phosphorylated TH4pep. (A) Two-dimensional (2D) distribution in which the values of the helical content measured for the region between D784 and Q792 are plotted against the distance between the carbon atom in the guanidinium group of R790 (atom CZ) and an oxygen atom of the phosphate of pS785 (atom O1P). Representative structures of the peptide extracted from the MD simulation are reported as cartoon. The D784, pS785 and R790 are shown as sticks. (B) 2D distribution in which the values of the helical content measured between D784 and Q792 are plotted against the distance between the carbon CZ of R790 and the γ-carbon atom of the D784 (atom CG). (C) 2D distribution in which the values of the distance between the carbon CZ of R790 and the oxygen atom of the phosphate of pS785 (atom O1P) are plotted against the distance between R790 CZ and the γ-carbon of the D784 (atom CG). (D) 2D distribution in which the values of the distance between the backbone amide hydrogen (atom H^N) and the oxygen atom of the phosphate of pS785 (atom O1P) are plotted against the angle measured between backbone amide nitrogen (atom N), H^N and O1P of pS785. (E) 2D distribution in which the values of the helical content measured between D784 and Q792 are plotted against the distance between the backbone amide hydrogen of S788 (atom H^N) and the oxygen atom of the phosphate of pS785 (atom O1P). (F) 2D distribution in which the values of the distance between the backbone amide hydrogen of S788 (atom H^N) and the oxygen atom of the phosphate of pS785 (atom O1P) are plotted against the angle measured between backbone amide nitrogen (atom N) and H^N of S788 and O1P of pS785.

strong electronegative group. This is in line with the observation that the amide peaks of the phosphorylated residues are still downfield shifted in 8 M urea (Supplementary Fig. S3D–E).

3.6. Atomic-level description of the structural effects induced by phosphorylation

The experimental data suggested that the increase in helicity upon S785 phosphorylation is accompanied by changes in the chemical shift of the R790 side chain ($H^{\varepsilon}N^{\varepsilon}$), as well as a significant increase of the temperature coefficient for the S788 backbone amide, suggesting the formation of specific non-covalent interactions. It is, however, difficult to extract the atomic details of these transiently formed structures from the chemical shifts or other NMR data. Thus, to map out atomic-level conformations of TH4pep that could explain the observations and provide insight into the formation of such transient interactions we conducted a 1-microsecond long all-atom, explicit solvent MD simulation of the phosphorylated TH4pep. In these simulations we employed the recently optimized Amber03ws force field (i.e. the physical model used to describe the system in the simulation), which was designed to improve the accuracy in MD simulations of disordered proteins and peptides, with a particular focus on removing a bias towards forming compact structures that is present in many earlier force fields [44]. Although the force field used in our simulations is expected to provide a relatively accurate description of the conformations of flexible peptides, we note that reaching converged conformational ensembles remains a difficult task due to the large number of possible conformations [61]. We thus here focus mostly on using the simulations to provide a structural interpretation of the experimental findings.

The conformational ensemble of the disordered TH4pep in solution was found to be highly heterogeneous and adopt multiple different structures. Indeed, we observed structures with varying amounts of helical structure and location, as can be attested by the average structures of the main structural clusters (Fig. 3A). We observed structures in which the side chains of R790 and pS785 interacted, though such interactions could be formed in both helical and non-helical structures (Fig. 3A). Also, we observed helical structures in which pS785 and R790 interacted in different orientations, as well as other structures where the two residues were separated, or where R790 interacted with the side chain of D784. We also did not find any correlation between the interaction between D784 and R790 and the amount of helix (Fig. 3B). To evaluate the contribution of D784 to the electrostatic interactions between R790 and pS785, we also calculated a two-dimensional density plot including the distance between the side chains of D784 and the R790 and the side chain of R790 and pS785 (Fig. 3C). The results show that when either D784 or pS785 forms a short, ionic interaction (distances ~5 Å), the other distance is longer (~8–10 Å; Fig. 3C). When neither distance is short, the two distances are correlated because of the inherent proximity of D784 and pS785.

The up-field shifted position of the amide H^N NMR peak of phosphorylated residues has commonly been attributed to intra-residual H-bond formation between backbone H^N and the phosphoryl-group [58,59]. We therefore analysed the presence of a hydrogen bond from the pS785 side chain to the backbone amides of pS785 (Fig. 3D), S788 (Fig. 3E–F) or R790 (Supplementary Fig. S5). To help define any such hydrogen bonds, we calculated the distance between acceptor and hydrogen atoms and the angle formed by the donor-hydrogen-acceptor (Fig. 3D–F). Consistently in the simulations, we did not observe any stable interaction between the phosphate moiety and the backbone amide of pS785, in agreement with the absence of a change in the temperature coefficient of S785 upon phosphorylation (Fig. 2D). However, we also did not observe any highly populated hydrogen bond between pS785 and the S788 backbone amide (Fig. 3E–F) although the phospho-group of pS785 was at a distance of approximately 4 Å in most of the MD frames, supporting also that the helix is only partially populated. This suggests that the large increase in the temperature coefficient of the backbone amide of S788 upon S785 phosphorylation is either due to a weak hydrogen bond not observed in the simulations, a hydrogen bond with a different side chain in close proximity, or that not a hydrogen bond but the close proximity of pS785 and R790 and the salt bridge formed between them (Fig. 3A) is responsible for the increased temperature coefficient, similarly to what is seen for ring current effects [60,62].

We also performed a one-microsecond long MD simulation of the unphosphorylated peptide (Supplementary Fig. S6). Again, as reaching converged distributions is computationally difficult, we decided not to perform a detailed, quantitative comparison of the two systems, in particular since our analyses pertain mostly to the interactions with and contributions from the phosphoryl group (Fig. 3C–F). In line with experiment, we do, however, observe a decreased helical propensity in our simulations of the unphosphorylated variant (Supplementary Fig. S6A, Fig. 3A), and also note that there was no correlation between helical formation and an electrostatic interaction between R790 and D784 also in absence of phosphorylation (Supplementary Fig. S6B, Fig. 3B).

Together our MD simulations of both the phosphorylated and unphosphorylated TH4pep allowed us to provide an atom-level model of the heterogeneity of the peptide, which showed no remarkable contribution by the D784 in stabilizing helical formation and, in the case of the phosphorylated peptide, also no stable intra-residue hydrogen bond between the phosphoryl-group and the backbone amide of the phospho-residue. The simulation of phosphorylated TH4pep instead revealed a propensity to form electrostatic interactions between the phosphate-group of pS785 and the side chain of R790.

3.7. Is [S/T]-P-{3}-[R/K] a common motif for helix stabilisation upon phosphorylation in other disordered regions?

The motif identified in NHE1 may constitute a new short linear motif (SLiM) [63] for phosphorylation-induced helix stabilisation with the core required amino acid sequence [S/T]-P-{3}-[R/K]. To investigate whether this motif is found in other IDPs we performed motif enrichment analysis using the D^2P^2 database [48]. In particular, we asked the question whether this motif was enriched in known phosphorylation sites in IDRs compared to IDRs with no known phosphorylation. Thus, we first extracted a background set that consisted of IDRs containing as a minimum a single Ser or Thr residue, but without any previously identified Ser/Thr phosphorylation sites, according to PhosphoSite [64,65] (in total 164,750 regions, 56,007 unique UniProt IDs). The query set of sequences was IDRs known to have at least one

Table 2
Motif enrichment for the six phosphorylation-sites in NHE1. The query motif represents the structural motif, whereas the known phosphor-motifs 1–4 are non-structural.

	Motif	AME Enrichment Adjusted *p*-value (uniform/individual frequencies of {3})	R Chi^2 test Adjusted *p*-value	R Odds ratio
Query motif (S693, S785)	[S/T]-P-x{3}-[R/K]	1.8×10^{-27}/6.9×10^{-13}	0.0	3.1
Known phospho-motif 1 (S723)	A-S-P-x{3}-E	1.3×10^{-9}/0.0013	1.23e-18	2.7
Known phospho-motif 2 (S771)	S-S-P-x{3}-D	0.006/0.0002	9.1e-13	2.1
Known phospho-motif 3 (S726)	Q-S-P-x{3}-D	0.002/not sig.	5.7e-09	2.8
Known phospho-motif 4 (T779)	F-T-P-x{3}-D	Not sig./not sig.	0.67	0.5

phosphorylated serine or threonine (in total 21,653 regions, 10,268 unique UniProt IDs). We enriched the query set of sequences for the presence of the motif [S/T]-P-{3}-[R/K], which we designed on the basis of the structural studies of the TH4pep constructs. As a comparison, we also performed enrichment analysis using four known NHE1 phosphorylation motifs where phosphorylation did not induce additional helical structure (termed non-structural sites, see above), these were A-S-P-{3}-E, S-S-P-{3}-D, Q-S-P-{3}-D and F-T-P-{3}-D (see also Fig. 1H). We used two different methods for motif search (see Materials and methods), i.e. Analysis of Motif Enrichment (AME) and an in-house developed methodology that relies on χ^2 tests and odds ratios (ORs).

In the enrichment analysis, we tested the effect of setting the distribution of the 3 "X" residues between P and R/K to either uniform or variable (to the values of the residue occurrence observed in the background) and found comparable results with both methods. Further, although we found that both the query motif and two of the three "non-structural" sites were enriched in phosphorylated IDRs, we observed that the adjusted *p*-values were substantially higher in the "non-structural" sites compared to the query motif. Further, the non-structural site F-T-P-{3}-D was not enriched in the query regions compared to background regions. The methodology implemented using R, i.e. χ^2 tests and ORs, revealed the same trend. All motifs, associated *p*-values and odds-ratios are reported in Table 2.

The substrates of ERK2 have a requirement for Pro in the i + 1 position. Prolines are very common in IDP/IDRs and are frequently observed in the flanking regions of transient structural elements. Pro is known as a cap residue [66], and MD simulations suggest a helix promoting effect for positions at the helical N-terminus, and a terminating effect at the C-terminus [67]. Thus, the proline in the query motif is required for the recognition by the kinase as a substrate. In addition, it may be important for defining the starting point of the helix.

Collectively the results of motif enrichment indicate that the novel query phospho-motif is enriched in disordered protein regions with phosphorylated Ser/Thr compared to disordered regions without this post-translation modification. While we cannot rule out that this observation is in part caused by preferences by the cellular kinases, the results suggest that this novel motif could have a general role as a conformational switch in IDRs. Intriguingly, perhaps the motif may act both to recruit the kinase and to confer a phosphorylation-induced conformational change. A full list of proteins (UniProt ID, gene name, organism) containing the [S/T]-P-{3}-[R/K] motif, is found in the Supporting information File S7. Based on these we created a sequence logo to visualize the motif that we discovered (Fig. 4). We note that the observed sequence profile reflects both the input sequence in our search as well as the distribution of amino acids surrounding the phosphorylation sites, and that further experiments are required to disentangle any potential preferences at the {3} positions.

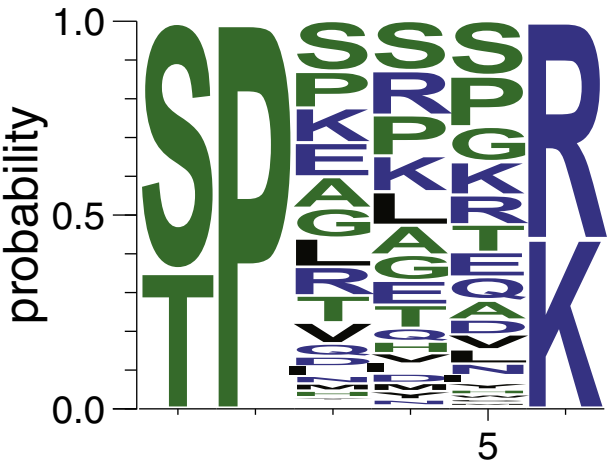


Fig. 4.

4. Discussion

4.1. Random coil chemical shifts of the phosphorylated state

Extraction of local structure from chemical shifts has been possible for almost 50 years [68] and the random coil chemical shift databases and peptide-derived libraries continue to improve both in accuracy and in correction factors for sequence and sample conditions like temperature and pH [19,21,22,69]. Critically, however, databases and libraries that incorporate the chemical effects of phosphorylations and other types of PTMs on the chemical shifts are currently not available. At present, the best possible way to extract the inducible effect of a phosphorylation, other PTMs, or non-natural amino acids is to assign a modified protein in the "fully unfolded state" and use this as internal random coil reference. We have shown here that this method is equally good, if not better, than database or library based methods to detect the position and amplitude of transient structure and additionally it differentially detects the effect of modification on the random coil state and potential structural modulation. The extraction of the structural effect by this method requires another round of assignment, and homogeneous phosphorylations of all sites as well as identical sample conditions for all states. Thus, until many more phosphorylated proteins have been assigned and their shifts and structures have become available, peptide derived random coil remain an efficient and accurate approach to examine the locally and globally induced structural changes of these modifications. With the internally referenced SCS analysis applied here, one can assess the local structural effect of individual phosphorylation sites at the residual level, and identify phosphorylation-induced structural modulations, as shown for the helix stabilizing effect of S785 phosphorylation on NHE1. This method further allows for distinguishing the change of chemical environment due to the modification from modulation of secondary structure by the modification. We expect this approach to be as useful for the determination of reference chemical shift values for other protein PTMs as it is for phosphorylation.

4.2. NHE1 phosphorylation – effects on secondary structure

In NHE1, TH1 and TH4 each have a phosphorylation site at their N-terminus, the phosphorylation of which leads to increased helix

population. Helix stabilisation is commonly attributed to hydrogen-bond formation between a capping residue with a negatively charged/phosphorylated side chain and one of the four unsatisfied backbone NH donors [70] as well as to favourable charge interaction with the helix dipole. While such capping may explain the increased helicity for TH1 and TH4, it does not explain the absence of such an effect of S726 phosphorylation for TH2 even though S726 is not as close to TH2 as S693 and S785 are to TH1 and TH4, respectively. Stabilisation can in addition to hydrogen bonding be explained by electrostatic interactions (either directly via salt bridges or via longer range interactions), i.e. the electrostatic attraction of the doubly negatively charged phosphate group to adjacent positively charged side chains. An electrostatic interaction is also less dependent on the geometry compared to a hydrogen bond. Positively charged residues close to S693 are H689 (i − 4), K690 (i − 3), and R698 (i + 5), and close to S785 is R790 (i + 5). Furthermore, both S693 and S785 have an Asp as the immediate neighbour (i − 1), which may contribute to the structural effect by electrostatic repulsion of the phosphate group. However, neither the experimental data nor simulations support a role of the Asp for the structural effect of phosphorylation. S771 has two N-terminal close positive charges, R765 (i − 6) and K767 (i − 4), but also an E768 (i − 3), all located at the C-terminus of TH3. Interestingly, none of the other phosphorylation sites have positive charges in their close proximity, and thus, it is likely that the N-terminal stabilisation of TH1 and TH4 is caused by electrostatic attraction as well as potential hydrogen bonds, as suggested by the MD simulations and supported by spectroscopic data.

4.3. Position dependency of structural effects induced by phosphorylation

The observation of a tuneable structure in NHE1cdt that depends on a specific phosphorylation, together with subsequent bioinformatics investigations of phosphorylation sites in IDPs, allowed us to identify a SLiM with the sequence [S/T]-P-{3}-[R/K], where [S/T] is the phosphorylation site. Although the idea of structure modulation by phosphorylation is not novel, the determinants for modulation of local structure by phosphorylation are poorly understood. Thus, examples of almost all possible scenarios are reported in the literature where the effect appears to depend on the position of the phosphorylation site within the helix as well as on the environment. Phosphorylations of the proline-rich IDR of Tau have distinct effects depending on the exact site, i.e. stabilisation of a short α-helix by forming an N-cap at one position, yet no effect on local structure at the other phosphorylation site [71]; very much similar to what is observed here for NHE1cdt. However, multiple phosphorylations in the same region in Tau were previously shown by CD and NMR spectroscopy to induce polyproline II helices [72]. Thus, the local sequence as well as the number of phosphorylation sites influences the structural outcome. However, typically, phosphorylation in the N-terminal of a helix results in stabilisation whereas a C-terminal phosphorylation results in destabilisation [9].

4.3.1. Helix stabilisation effects

Commonly, N-terminal helix stabilisation is attributed to hydrogen-bond formation between a capping residue with a negatively charged or polar side chain and one of the four unsatisfied backbone NH donors [70]. At physiological pH each phosphorylation adds two negative charges to the protein. Accordingly, a phosphorylation-induced N-terminal stabilisation was predicted by simulations [10]. However, when a phosphorylation is stabilizing, a salt bridge is commonly formed between the phosphoryl-group and some Arg/Lys residues, seen both in folded protein complexes and in disordered proteins as described with specific examples further below. This suggests that the phosphoryl group both forms interactions with one of these amides and forms a salt bridge to the Arg/Lys side chain, although only the latter interaction is well documented. Depending on the protein, phosphorylation can have dramatic structural effects. In the cyclin-dependent kinase CDK2, three Arg side chains make charge-stabilized hydrogen bonds to the three oxygen atoms of the phosphate of pThr160, leading to large global structural rearrangements and kinase activation [55]. In the enzyme phenylalanine hydroxylase, phosphorylation in the N-terminal IDR also induces a salt bridge to an Arg in position i − 3 leading to a local stabilisation but no global rearrangements [54]. In the phosphorylation domain of smooth muscle myosin, electron paramagnetic resonance data and MD simulations suggested a disorder-to-order transition for the phosphorylation domain mediated by a salt bridge between pS19 and R16 [56,73,74]. Furthermore, in a pentamer peptide derived from the cyclic-AMP dependent protein kinase, phosphorylation led to formations of a hydrogen bond to an Arg in i − 3, accompanied by a local rearrangement [11]. Others have shown that stabilisation is achieved via a hydrogen bond network plus electrostatic interactions of the phosphate with adjacent positive charged residues not only in the N-cap position but also in other structural entities [53–55,75]. Similarly, a helix-stabilizing effect was also found for phosphorylation of S65 of eIF4E, in this case at the C-terminal end of a helical region [76]. Interestingly, for the disordered C-terminal domain of connexin43, the effect of phosphorylation on helicity was only observed in a membrane bound state [12], highlighting the relevance of the local environment. Thus, although capping to unpaired amide donors appears to dominate in helices, a positively charged residue at position i − 3 governs structure induction. Here we show that also a positively charged residue in position i + 5 should be considered and that this potentially is accompanied by capping of one unpaired backbone amide donor which is not the phosphorylated residue's own amide.

4.3.2. Helix destabilizing effects

Phosphorylation can also have a destabilizing effect on the structure, as shown for phospholamban where a salt bridge occurs within the C-terminal part of a helix, with a concomitant loss of helicity [77,78]. S16 phosphorylation destabilizes the local structure elements [77,79], and C-terminal phosphorylation of a helix in 4E-BP1 (a regulatory factor for the translation initiation factor elF4E) similarly destabilizes local structure, resulting in failure to bind eIF4E [76].

Taken together, the effects of phosphorylation depend strongly on the position and context of the phosphorylated residue and any neighbouring residues with which interactions can be formed. Thus, the tuneable motif determined here shows distinctly that a pSer, Arg pair at i, i + 5 in a helical N-cap position is stabilizing. The diverse and seemingly uncorrelated observations made in the literature may originate from the presence of several independent SLiMs, some of which are switchable and some of which are not. Here we have defined one SLiM, which may be able to explain previous observations but more importantly may spur an interest in defining additional such motifs.

4.4. Switch–like behaviour

Modulation of a structural element that is required for an interaction provides a general mechanism of functional *on-off* switches. Accordingly, an increase in helical propensity within an IDP was recently suggested to accelerate ligand-binding [6]. Independent of the presence of charged groups in the surrounding sequence, Ser/Thr phosphorylation was suggested to modulate the preferred backbone dihedral angles [15] priming the potential for structure modulation. In a peptide study it was examined whether a Lys in i, i + 3 and i, i + 4 relative to a pSer within a helix could have the same effect and it was concluded that the presence of the Lys allowed salt bridge formation and modulated the effect of phosphorylation. Thus in the absence of a Lys, phosphorylation could destabilize the helix (> 1 kcal·mol^{-1}) where as in its presence, a phosphorylation was stabilizing (< -2 kcal·mol^{-1}) [80].

Importantly, despite the wide use of phospho-mimetic mutations (Ser/Thr to Asp/Glu) [81], computational and experimental studies have shown that phospho-mimetic mutations are often insufficient to

fully simulate the effect of phosphorylation [82–84]. For example, phosphorylation of the disordered activation domain of the glucocorticoid receptor by p38 MAPK induces secondary and tertiary structure formation that facilitates interaction with co-regulators. In contrast, a Ser to Glu mutation in the same location elicited only a modest increase in helicity [83]. These effects can perhaps be understood in the frame of the dualistic properties of the phosphoryl group; it needs to act as a cap, but at the same time be involved in a salt bridge formation to strategically positioned Arg/Lys-side chains. Thus, we suggest this combined effect to be termed as *phospho-cap*.

The local sequence of the NHE1 TH4 forms a lowly populated helix (5.6%) that gets stabilized to approximately 16% upon phosphorylation, due to the combined effects of electrostatic interaction with the Arg in i + 5 and potentially a weak hydrogen bond formation with the amide in i + 3. This represents a small, but significant change in helicity, corresponding to an increase in the stability of the helix by ~ 0.5 kcal mol^{-1}, leading to a shift of the dynamic ensemble typical for disordered proteins. Additional studies are needed in order to identify factors that can potentially further modulate the helix, and the effect of helix formation on e.g. the interaction with other proteins. This however requires the identification of binding partners of this particular site and its functional role in vivo. A number of identified binding partners for NHE1 were shown to interact with the most distal part of the C-tail [32], i.e. including TH4 (P786-L795), and could thus be potential candidates for phosphorylation-dependent binding switching. Interestingly, the interaction of NHE1 with carbonic anhydrase II, which has been proposed to interact within the last C-terminal 130 residues of NHE1, was regulated by an NHE1 phosphorylation proximal to the binding region [85].

The *phospho-cap* motif was identified in a large number of other proteins all of diverse functions and structures (Supplementary information File S7). One interesting observation is that many of these are themselves kinases or are kinase-anchoring proteins or receptor-associated proteins. This suggests that phosphorylation inducible structure formation is part of several phosphorylation cascades and play a role in signalling, e.g. the mitogen activated protein kinase kinases, the calcium/calmodulin dependent kinases, and the protein kinase C binding proteins. Remarkably, a similar motif has been described for cyclin dependent kinases, which have preference for the [S/T]P-{1,2}-[R/K] motif [86,87], one or two residue(s) shorter than the phospho-cap motif described here, although some show preference for the long motif identical to the phospho-cap [86]. However, it has been shown that cyclin dependent kinases phosphorylates BRCA2 at S3291, which is located in a longer motif, SPAAQK$_{3296}$ identical to the one discovered here. Phosphorylation of S3291 subsequently blocks the interaction with RAD51, an essential recombination protein [88], but whether this involves stabilisation of helical structure via formation of a phospho-cap, remains to be verified. Another implication of these similarities is that the enrichment of the motif we observe may in fact be due to a dominance of kinase preference sites, and suggest a possible dual role in recruiting the kinase as well as acting as the phosphor-induced-structure sensor. Given that numerous mitogen-activated protein kinases (MAPKs) and their regulators are part of the NHE1 interactome [89], it is interesting that from this list we identified several of these, such as SLK, TAO3, and MAP3K4. Together with the fact that the list also contains several other ion transport proteins, this opens the possibility that interactions between such proteins are regulated by phosphorylation-induced altered helix propensity of one or both partners.

5. Conclusions

The effect of phosphorylation depends on several factors, i.e. the presence and type of transient secondary structure elements, the relative position of the phosphorylation site to preformed structural elements, the adjacent sequence including the consensus phosphorylation motif of the kinase, and the local environment of the protein, e.g. at the membrane. Further, the effect is dependent on the type of phosphorylated residue, i.e. Ser, Thr, or Tyr. Thus, the concerted contributions of all those factors determine the net modulation in local structure, but as of yet, the effects cannot be bioinformatically predicted. Using a combination of experiment, simulation and bioinformatics we have identified a tuneable motif that allows stabilisation of helical structure upon phosphorylation. We believe this motif is the first of its kind that has been discovered, but also that it is likely that several others exist that may underlie the diverse effects reported in the literature. The identification and validation of more motifs could have a large impact on our understanding of the molecular bases underlying phosphorylation-mediated cellular regulation, will help to develop predictors of structural effects by phosphorylation, and may be extremely useful in designing switches into proteins. Ultimately, such knowledge will provide us with the predictive power needed to understand why some mutations involving phosphorylation sites may give rise to disease.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.cellsig.2017.05.015.

Author contribution

R.H.A., K.L.-L., and B.B.K. conceived the idea and designed the outline of the work with input from E.P. All protein purification, CD and NMR analyses were done by R.H.A., while T.T. and E.P. performed the bioinformatics analyses and M.L., E.P. and K.L.-L. carried out molecular modeling and subsequent analyses. The phosphorylation analyses were done by R.H.A. with S.F.P. and B.B.K. R.H.A. and B.B.K. wrote the manuscript with input from all authors, mainly S.F.P., E.P. and K.L.-L.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The result of this research has been achieved using the ISCRA-CINECA HPC Grant ALLO-PCM (HP10CWP9KW) on Galileo HPC Cluster to E.P. E.P.'s group is supported by a starting grant from the Danish Cancer Society Research Center (DCRC), Innovation Fund Denmark and it is part of the Center of Excellence in Autophagy, Recycling and Disease (CARD), funded by the Danish National Research Foundation. This work was further supported by grants from the Danish National Research Council (B.B.K.: 12-128803 and B.B.K./S.F.P.: DFF – 4181-00344) and a Hallas-Møller stipend from the Novo Nordisk Foundation to K.L.-L.

References

[1] P.E. Wright, H.J. Dyson, Linking folding and binding, Curr. Opin. Struct. Biol. 19 (1) (2009) 31–38, http://dx.doi.org/10.1016/j.sbi.2008.12.003.
[2] H.J. Dyson, P.E. Wright, Coupling of folding and binding for unstructured proteins, Curr. Opin. Struct. Biol. 12 (1) (2002) 54–60, http://dx.doi.org/10.1016/S0959-440X(02)00289-0.
[3] J.M. Rogers, V. Oleinikovas, S.L. Shammas, C.T. Wong, D. De Sancho, C.M. Baker, J. Clarke, Interplay between partner and ligand facilitates the folding and binding of an intrinsically disordered protein, Proc. Natl. Acad. Sci. U. S. A. 111 (43) (2014) 15420–15425, http://dx.doi.org/10.1073/pnas.1409122111.
[4] J.M. Rogers, A. Steward, J. Clarke, Folding and binding of an intrinsically disordered protein: fast, but not 'diffusion-limited", J. Am. Chem. Soc. 135 (4) (2013) 1415–1422, http://dx.doi.org/10.1021/ja309527h.
[5] K. Sugase, H.J. Dyson, P.E. Wright, Mechanism of coupled folding and binding of an intrinsically disordered protein, Nature 447 (7147) (2007) 1021–1025, http://dx.doi.org/10.1038/nature05858.
[6] V. Iesmantavicius, J. Dogan, P. Jemth, K. Teilum, M. Kjaergaard, Helical propensity in an intrinsically disordered protein accelerates ligand binding, Angew. Chem. Int. Ed. Eng. 53 (6) (2014) 1548–1551, http://dx.doi.org/10.1002/anie.201307712.
[7] A. Bah, R.M. Vernon, Z. Siddiqui, M. Krzeminski, R. Muhandiram, C. Zhao, N. Sonenberg, L.E. Kay, J.D. Forman-Kay, Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch, Nature 519 (7541) (2015) 106–109, http://dx.doi.org/10.1038/nature13999.
[8] E.W. Martin, A.S. Holehouse, C.R. Grace, A. Hughes, R.V. Pappu, T. Mittag,

Sequence determinants of the conformational properties of an intrinsically disordered protein prior to and upon multisite phosphorylation, J. Am. Chem. Soc. 138 (47) (2016) 15323–15335, http://dx.doi.org/10.1021/jacs.6b10272.
[9] C.D. Andrew, J. Warwicker, G.R. Jones, A.J. Doig, Effect of phosphorylation on alpha-helix stability as a function of position, Biochemistry 41 (6) (2002) 1897–1905, http://dx.doi.org/10.1021/bi0113216.
[10] J.L. Smart, J.A. McCammon, Phosphorylation stabilizes the N-termini of alpha-helices, Biopolymers 49 (3) (1999) 225–233, http://dx.doi.org/10.1002/(SICI)1097-0282(199903)49:3<225::AID-BIP4>3.0.CO;2-B.
[11] Y.J. Kang, L.M. Zuo, S.Z. Luo, Hydrogen-bonding interactions induced by phosphorylation influence the local conformation of phosphopeptides, Int. J. Pept. Res. Ther. 16 (2) (2010) 87–93, http://dx.doi.org/10.1007/s10989-010-9207-y.
[12] R. Grosely, J.L. Kopanic, S. Nabors, F. Kieken, G. Spagnol, M. Al-Mugotir, S. Zach, P.L. Sorgen, Effects of phosphorylation on the structure and backbone dynamics of the intrinsically disordered connexin43 C-terminal domain, J. Biol. Chem. 288 (34) (2013) 24857–24870, http://dx.doi.org/10.1074/jbc.M113.454389.
[13] G. Desjardins, C.A. Meeker, N. Bhachech, S.L. Currie, M. Okon, B.J. Graves, L.P. McIntosh, Synergy of aromatic residues and phosphoserines within the intrinsically disordered DNA-binding inhibitory elements of the Ets-1 transcription factor, Proc. Natl. Acad. Sci. U. S. A. 111 (30) (2014) 11019–11024, http://dx.doi.org/10.1073/pnas.1401891111.
[14] P.J. Roach, Multisite and hierarchal protein phosphorylation, J. Biol. Chem. 266 (22) (1991) 14139–14142.
[15] A. Tholey, A. Lindemann, V. Kinzel, J. Reed, Direct effects of phosphorylation on the preferred backbone conformation of peptides: a nuclear magnetic resonance study, Biophys. J. 76 (1 Pt 1) (1999) 76–87, http://dx.doi.org/10.1016/S0006-3495(99)77179-1.
[16] D.S. Wishart, B.D. Sykes, The 13C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data, J. Biomol. NMR 4 (2) (1994) 171–180, http://dx.doi.org/10.1007/bf00175245.
[17] D.S. Wishart, B.D. Sykes, F.M. Richards, The chemical-shift index — a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, Biochemistry 31 (6) (1992) 1647–1651, http://dx.doi.org/10.1021/bi00121a010.
[18] D.S. Wishart, C.G. Bigam, A. Holm, R.S. Hodges, B.D. Sykes, ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects, J. Biomol. NMR 5 (1) (1995) 67–81, http://dx.doi.org/10.1007/bf00227471.
[19] M. Kjaergaard, S. Brander, F.M. Poulsen, Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH, J. Biomol. NMR 49 (2) (2011) 139–149, http://dx.doi.org/10.1007/s10858-011-9472-x.
[20] J.A. Marsh, V.K. Singh, Z. Jia, J.D. Forman-Kay, Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation, Protein Sci. 15 (12) (2006) 2795–2804, http://dx.doi.org/10.1110/ps.062465306.
[21] M. Kjaergaard, F.M. Poulsen, Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution, J. Biomol. NMR 50 (2) (2011) 157–165, http://dx.doi.org/10.1007/s10858-011-9508-2.
[22] K. Tamiola, B. Acar, F.A. Mulder, Sequence-specific random coil chemical shifts of intrinsically disordered proteins, J. Am. Chem. Soc. 132 (51) (2010) 18000–18003, http://dx.doi.org/10.1021/ja105656t.
[23] E.A. Bienkiewicz, K.J. Lumb, Random-coil chemical shifts of phosphorylated amino acids, J. Biomol. NMR 15 (3) (1999) 203–206, http://dx.doi.org/10.1023/a:1008375029746.
[24] K. Modig, V.W. Jurgensen, K. Lindorff-Larsen, W. Fieber, H.G. Bohr, F.M. Poulsen, Detection of initiation sites in protein folding of the four helix bundle ACBP by chemical shift analysis, FEBS Lett. 581 (25) (2007) 4965–4971, http://dx.doi.org/10.1016/j.febslet.2007.09.027.
[25] W.K. Lim, J. Rosgen, S.W. Englander, Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group, Proc. Natl. Acad. Sci. U. S. A. 106 (8) (2009) 2595–2600, http://dx.doi.org/10.1073/pnas.0812588106.
[26] W.A. Elam, T.P. Schrank, A.J. Campagnolo, V.J. Hilser, Temperature and urea have opposing impacts on polyproline II conformational bias, Biochemistry 52 (5) (2013) 949–958, http://dx.doi.org/10.1021/bi301435p.
[27] S.J. Whittington, B.W. Chellgren, V.M. Hermann, T.P. Creamer, Urea promotes polyproline II helix formation: implications for protein denatured states, Biochemistry 44 (16) (2005) 6269–6275, http://dx.doi.org/10.1021/bi050124u.
[28] K. Modig, F.M. Poulsen, Model-independent interpretation of NMR relaxation data for unfolded proteins: the acid-denatured state of ACBP, J. Biomol. NMR 42 (3) (2008) 163–177, http://dx.doi.org/10.1007/s10858-008-9280-0.
[29] H.I. Rosner, F.M. Poulsen, Residue-specific description of non-native transient structures in the ensemble of acid-denatured structures of the all-beta protein c-src SH3, Biochemistry 49 (15) (2010) 3246–3253, http://dx.doi.org/10.1021/bi902125j.
[30] G.W. Haxholm, L.F. Nikolajsen, J.G. Olsen, J. Fredsted, F.H. Larsen, V. Goffin, S.F. Pedersen, A.J. Brooks, M.J. Waters, B.B. Kragelund, Intrinsically disordered cytoplasmic domains of two cytokine receptors mediate conserved interactions with membranes, Biochem. J. 468 (3) (2015) 495–506, http://dx.doi.org/10.1042/BJ20141243.
[31] R. Hendus-Altenburger, E. Pedraz-Cuesta, C.W. Olesen, E. Papaleo, J.A. Schnell, J.T. Hopper, C.V. Robinson, S.F. Pedersen, B.B. Kragelund, The human Na(+)/H(+) exchanger 1 is a membrane scaffold protein for extracellular signal-regulated kinase 2, BMC Biol. 14 (2016) 31, http://dx.doi.org/10.1186/s12915-016-0252-7.
[32] R. Hendus-Altenburger, B.B. Kragelund, S.F. Pedersen, Structural dynamics and regulation of the mammalian SLC9A family of Na(+)/H(+) exchangers, Curr. Top. Membr. 73 (2014) 69–148, http://dx.doi.org/10.1016/B978-0-12-800223-0.00002-5.
[33] E. Boedtkjer, L. Bunch, S.F. Pedersen, Physiology, pharmacology and pathophysiology of the pH regulatory transport proteins NHE1 and NBCn1: similarities, differences, and implications for cancer therapy, Curr. Pharm. Des. 18 (10) (2012) 1345–1371, http://dx.doi.org/10.2174/138161212799504830.
[34] A.B. Norholm, R. Hendus-Altenburger, G. Bjerre, M. Kjaergaard, S.F. Pedersen, B.B. Kragelund, The intracellular distal tail of the Na^+/H^+ exchanger NHE1 is intrinsically disordered: implications for NHE1 trafficking, Biochemistry 50 (17) (2011) 3469–3480, http://dx.doi.org/10.1021/bi1019989.
[35] S.F. Pedersen, B.V. Darborg, M.L. Rentsch, M. Rasmussen, Regulation of mitogen-activated protein kinase pathways by the plasma membrane Na^+/H^+ exchanger, NHE1, Arch. Biochem. Biophys. 462 (2) (2007) 195–201, http://dx.doi.org/10.1016/j.abb.2006.12.001.
[36] S.F. Pedersen, B.V. Darborg, M. Rasmussen, J. Nylandsted, E.K. Hoffmann, The Na^+/H^+ exchanger, NHE1, differentially regulates mitogen-activated protein kinase subfamilies after osmotic shrinkage in Ehrlich Lettre ascites cells, Cell. Physiol. Biochem. 20 (6) (2007) 735–750, http://dx.doi.org/10.1159/000110434.
[37] F. Delaglio, S. Grzesiek, G. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J. Biomol. NMR 6 (3) (1995) 277–293, http://dx.doi.org/10.1007/BF00197809.
[38] W.F. Vranken, W. Boucher, T.J. Stevens, R.H. Fogh, A. Pajon, M. Llinas, E.L. Ulrich, J.L. Markley, J. Ionides, E.D. Laue, The CCPN data model for NMR spectroscopy: development of a software pipeline, Proteins 59 (4) (2005) 687–696, http://dx.doi.org/10.1002/prot.20449.
[39] V. Munoz, L. Serrano, Elucidating the folding problem of helical peptides using empirical parameters, Nat. Struct. Biol. 1 (6) (1994) 399–409, http://dx.doi.org/10.1038/nsb0694-399.
[40] D.V. Fedyukina, S. Rajagopalan, A. Sekhar, E.C. Fulmer, Y.J. Eun, S. Cavagnero, Contribution of long-range interactions to the secondary structure of an unfolded globin, Biophys. J. 99 (5) (2010) L37–L39, http://dx.doi.org/10.1016/j.bpj.2010.06.038.
[41] A. Irback, S. Mohanty, PROFASI: a Monte Carlo simulation package for protein folding and aggregation, J. Comput. Chem. 27 (13) (2006) 1548–1555, http://dx.doi.org/10.1002/jcc.20452.
[42] A. Irback, S. Mohanty, Folding thermodynamics of peptides, Biophys. J. 88 (3) (2005) 1560–1569, http://dx.doi.org/10.1529/biophysj.104.050427.
[43] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, J. Chem. Theory Comput. 4 (3) (2008) 435–447, http://dx.doi.org/10.1021/ct700301q.
[44] R.B. Best, W. Zheng, J. Mittal, Balanced protein-water interactions improve properties of disordered proteins and non-specific protein association, J. Chem. Theory Comput. 10 (11) (2014) 5113–5124, http://dx.doi.org/10.1021/ct500569b.
[45] G.A. Khoury, J.P. Thompson, J. Smadbeck, C.A. Kieslich, C.A. Floudas, Forcefield_PTM: Ab initio charge and AMBER forcefield parameters for frequently occurring post-translational modifications, J. Chem. Theory Comput. 9 (12) (2013) 5653–5674, http://dx.doi.org/10.1021/ct400556v.
[46] I.Y. Torshin, I.T. Weber, R.W. Harrison, Geometric criteria of hydrogen bonds in proteins and identification of 'bifurcated' hydrogen bonds, Protein Eng. 15 (5) (2002) 359–363, http://dx.doi.org/10.1093/protein/15.5.359.
[47] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (12) (1983) 2577–2637, http://dx.doi.org/10.1002/bip.360221211.
[48] M.E. Oates, P. Romero, T. Ishida, M. Ghalwash, M.J. Mizianty, B. Xue, Z. Dosztanyi, V.N. Uversky, Z. Obradovic, L. Kurgan, A.K. Dunker, J. Gough, D(2)P(2): database of disordered protein predictions, Nucleic Acids Res. 41 (Database issue) (2013) D508–D516, http://dx.doi.org/10.1093/nar/gks1226.
[49] N. Xiao, D.S. Cao, M.F. Zhu, Q.S. Xu, protr/ProtrWeb: R package and web server for generating various numerical representation schemes of protein sequences, Bioinformatics 31 (11) (2015) 1857–1859, http://dx.doi.org/10.1093/bioinformatics/btv042.
[50] R.K. Saiki, S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn, H.A. Erlich, N. Arnheim, Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia, Science 230 (4732) (1985) 1350–1354, http://dx.doi.org/10.1126/science.2999980.
[51] F.A. Buske, M. Boden, D.C. Bauer, T.L. Bailey, Assigning roles to DNA regulatory motifs using comparative genomics, Bioinformatics 26 (7) (2010) 860–866, http://dx.doi.org/10.1093/bioinformatics/btq049.
[52] T.L. Bailey, M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, J. Ren, W.W. Li, W.S. Noble, MEME SUITE: tools for motif discovery and searching, Nucleic Acids Res. 37 (Web Server issue) (2009) W202–W208, http://dx.doi.org/10.1093/nar/gkp335.
[53] Y.-J. Kang, L.-M. Zuo, S.-Z. Luo, Hydrogen-bonding interactions induced by phosphorylation influence the local conformation of phosphopeptides, Int. J. Pept. Res. Ther. 16 (2) (2010) 87–93, http://dx.doi.org/10.1007/s10989-010-9207-y.
[54] F.F. Miranda, M. Thorolfsson, K. Teigen, J.M. Sanchez-Ruiz, A. Martinez, Structural and stability effects of phosphorylation: localized structural changes in phenylalanine hydroxylase, Protein Sci. 13 (5) (2004) 1219–1226, http://dx.doi.org/10.1110/ps.03595904.
[55] A.A. Russo, P.D. Jeffrey, N.P. Pavletich, Structural basis of cyclin-dependent kinase activation by phosphorylation, Nat. Struct. Biol. 3 (8) (1996) 696–700, http://dx.doi.org/10.1038/nsb0896-696.
[56] L.M. Espinoza-Fonseca, D. Kast, D.D. Thomas, Thermodynamic and structural basis of phosphorylation-induced disorder-to-order transition in the regulatory light chain of smooth muscle myosin, J. Am. Chem. Soc. 130 (37) (2008) 12208–12209, http://dx.doi.org/10.1021/ja803143g.

[57] M.C. Manning, R.W. Woody, Theoretical CD studies of polypeptide helices: examination of important electronic and geometric factors, Biopolymers 31 (5) (1991) 569–586, http://dx.doi.org/10.1002/bip.360310511.
[58] J.T. Du, Y.M. Li, W. Wei, G.S. Wu, Y.F. Zhao, K. Kanazawa, T. Nemoto, H. Nakanishi, Low-barrier hydrogen bond between phosphate and the amide group in phosphopeptide, J. Am. Chem. Soc. 127 (47) (2005) 16350–16351, http://dx.doi.org/10.1021/ja054568p.
[59] T.A. Ramelot, L.K. Nicholson, Phosphorylation-induced structural changes in the amyloid precursor protein cytoplasmic tail detected by NMR, J. Mol. Biol. 307 (3) (2001) 871–884, http://dx.doi.org/10.1006/jmbi.2001.4535.
[60] T. Cierpicki, J. Otlewski, Amide proton temperature coefficients as hydrogen bond indicators in proteins, J. Biomol. NMR 21 (3) (2001) 249–261, http://dx.doi.org/10.1023/a:1012911329730.
[61] Z.A. Levine, J.E. Shea, Simulations of disordered proteins and systems with conformational heterogeneity, Curr. Opin. Struct. Biol. 43 (2017) 95–103, http://dx.doi.org/10.1016/j.sbi.2016.11.006.
[62] G. Merutka, H.J. Dyson, P.E. Wright, 'Random coil' 1H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG, J. Biomol. NMR 5 (1) (1995) 14–24, http://dx.doi.org/10.1007/bf00227466.
[63] N.E. Davey, N.J. Haslam, D.C. Shields, R.J. Edwards, SLiMSearch 2.0: biological context for short linear motifs in proteins, Nucleic Acids Res. 39 (Web Server issue) (2011) W56–W60, http://dx.doi.org/10.1093/nar/gkr402.
[64] P.V. Hornbeck, I. Chabra, J.M. Kornhauser, E. Skrzypek, B. Zhang, PhosphoSite: a bioinformatics resource dedicated to physiological protein phosphorylation, Proteomics 4 (6) (2004) 1551–1561, http://dx.doi.org/10.1002/pmic.200300772.
[65] P.V. Hornbeck, B. Zhang, B. Murray, J.M. Kornhauser, V. Latham, E. Skrzypek, PhosphoSitePlus, 2014: mutations, PTMs and recalibrations, Nucleic Acids Res. 43 (Database issue) (2015) D512–D520, http://dx.doi.org/10.1093/nar/gku1267.
[66] F.X. Theillet, L. Kalmár, P. Tompa, K.H. Han, P. Selenko, A.K. Dunker, G.W. Daughdrill, V.N. Uversky, The alphabet of intrinsic disorder: I. Act like a Pro: on the abundance and roles of proline residues in intrinsically disordered proteins, Intrinsically Disord. Proteins 1 (1) (2013) e24360, http://dx.doi.org/10.4161/idp.24360.
[67] C. Lee, L. Kalmar, B. Xue, P. Tompa, G.W. Daughdrill, V.N. Uversky, K.H. Han, Contribution of proline to the pre-structuring tendency of transient helical secondary structure elements in intrinsically disordered proteins, Biochim. Biophys. Acta 1840 (3) (2014) 993–1003, http://dx.doi.org/10.1016/j.bbagen.2013.10.042.
[68] J.L. Markley, D.H. Meadows, O. Jardetzky, Nuclear magnetic resonance studies of helix-coil transitions in polyamino acids, J. Mol. Biol. 27 (1) (1967) 25–40, http://dx.doi.org/10.1016/0022-2836(67)90349-x.
[69] S. Schwarzinger, G.J. Kroon, T.R. Foss, J. Chung, P.E. Wright, H.J. Dyson, Sequence-dependent correction of random coil NMR chemical shifts, J. Am. Chem. Soc. 123 (13) (2001) 2970–2978, http://dx.doi.org/10.1021/ja003760i.
[70] R. Aurora, G.D. Rose, Helix capping, Protein Sci. 7 (1) (1998) 21–38, http://dx.doi.org/10.1002/pro.5560070103.
[71] N. Sibille, I. Huvent, C. Fauquant, D. Verdegem, L. Amniai, A. Leroy, J.M. Wieruszeski, G. Lippens, I. Landrieu, Structural characterization by nuclear magnetic resonance of the impact of phosphorylation in the proline-rich region of the disordered Tau protein, Proteins 80 (2) (2012) 454–462, http://dx.doi.org/10.1002/prot.23210.
[72] A.A. Bielska, N.J. Zondlo, Hyperphosphorylation of tau induces local polyproline II helix, Biochemistry 45 (17) (2006) 5527–5537, http://dx.doi.org/10.1021/bi052662c.
[73] W.D. Nelson, S.E. Blakely, Y.E. Nesmelov, D.D. Thomas, Site-directed spin labeling reveals a conformational switch in the phosphorylation domain of smooth muscle myosin, Proc. Natl. Acad. Sci. U. S. A. 102 (11) (2005) 4000–4005, http://dx.doi.org/10.1073/pnas.0401664102.
[74] L.M. Espinoza-Fonseca, D. Kast, D.D. Thomas, Molecular dynamics simulations reveal a disorder-to-order transition on phosphorylation of smooth muscle myosin, Biophys. J. 93 (6) (2007) 2083–2090, http://dx.doi.org/10.1529/biophysj.106.095802.
[75] J. Li, D.J. Bigelow, T.C. Squier, Phosphorylation by cAMP-dependent protein kinase modulates the structural coupling between the transmembrane and cytosolic domains of phospholamban, Biochemistry 42 (36) (2003) 10674–10682, http://dx.doi.org/10.1021/bi034708c.
[76] S. Tait, K. Dutta, D. Cowburn, J. Warwicker, A.J. Doig, J.E. McCarthy, Local control of a disorder-order transition in 4E-BP1 underpins regulation of translation via eIF4E, Proc. Natl. Acad. Sci. U. S. A. 107 (41) (2010) 17627–17632, http://dx.doi.org/10.1073/pnas.1008242107.
[77] M.G. Paterlini, D.D. Thomas, The alpha-helical propensity of the cytoplasmic domain of phospholamban: a molecular dynamics simulation of the effect of phosphorylation and mutation, Biophys. J. 88 (5) (2005) 3243–3251, http://dx.doi.org/10.1529/biophysj.104.054460.
[78] R.J. Mortishire-Smith, S.M. Pitzenberger, C.J. Burke, C.R. Middaugh, V.M. Garsky, R.G. Johnson, Solution structure of the cytoplasmic domain of phospholamban: phosphorylation leads to a local perturbation in secondary structure, Biochemistry 34 (23) (1995) 7603–7613, http://dx.doi.org/10.1021/bi00023a006.
[79] E.E. Metcalfe, N.J. Traaseth, G. Veglia, Serine 16 phosphorylation induces an order-to-disorder transition in monomeric phospholamban, Biochemistry 44 (11) (2005) 4386–4396, http://dx.doi.org/10.1021/bi047571e.
[80] N. Errington, A.J. Doig, A phosphoserine-lysine salt bridge within an alpha-helical peptide, the strongest alpha-helix side-chain interaction measured to date, Biochemistry 44 (20) (2005) 7553–7558, http://dx.doi.org/10.1021/bi050297j.
[81] N.A. Sieracki, Y.A. Komarova, Studying Cell Signal Transduction With Biomimetic Point Mutations, Genetic Manipulation of DNA and Protein - Examples From Current Research, INTECH Open Access Publisher, 2013, pp. 381–392, http://dx.doi.org/10.5772/35029.
[82] E.S. Groban, A. Narayanan, M.P. Jacobson, Conformational changes in protein loops and helices induced by post-translational phosphorylation, PLoS Comput. Biol. 2 (4) (2006) e32, , http://dx.doi.org/10.1371/journal.pcbi.0020032.
[83] A.M. Garza, S.H. Khan, R. Kumar, Site-specific phosphorylation induces functionally active conformation in the intrinsically disordered N-terminal activation function (AF1) domain of the glucocorticoid receptor, Mol. Cell. Biol. 30 (1) (2010) 220–230, http://dx.doi.org/10.1128/MCB.00552-09.
[84] K.E. Paleologou, A. Oueslati, G. Shakked, C.C. Rospigliosi, H.Y. Kim, G.R. Lamberto, C.O. Fernandez, A. Schmid, F. Chegini, W.P. Gai, D. Chiappe, M. Moniatte, B.L. Schneider, P. Aebischer, D. Eliezer, M. Zweckstetter, E. Masliah, H.A. Lashuel, Phosphorylation at S87 is enhanced in synucleinopathies, inhibits alpha-synuclein oligomerization, and influences synuclein-membrane interactions, J. Neurosci. 30 (9) (2010) 3184–3198, http://dx.doi.org/10.1523/JNEUROSCI.5922-09.2010.
[85] X. Li, Y. Liu, B.V. Alvarez, J.R. Casey, L. Fliegel, A novel carbonic anhydrase II binding site regulates NHE1 activity, Biochemistry 45 (7) (2006) 2414–2424, http://dx.doi.org/10.1021/bi051132d.
[86] R. Hodeify, A. Tarcsafalvi, J. Megyesi, R.L. Safirstein, P.M. Price, Cdk2-dependent phosphorylation of p21 regulates the role of Cdk2 in cisplatin cytotoxicity, Am. J. Physiol. Ren. Physiol. 300 (5) (2011) F1171–F1179, http://dx.doi.org/10.1152/ajprenal.00507.2010.
[87] J.A. Lees, K.J. Buchkovich, D.R. Marshak, C.W. Anderson, E. Harlow, The retinoblastoma protein is phosphorylated on multiple sites by human Cdc2, EMBO J. 10 (13) (1991) 4279–4290.
[88] F. Esashi, N. Christ, J. Gannon, Y. Liu, T. Hunt, M. Jasin, S.C. West, CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair, Nature 434 (7033) (2005) 598–604, http://dx.doi.org/10.1038/nature03404.
[89] S. Bandyopadhyay, C.Y. Chiang, J. Srivastava, M. Gersten, S. White, R. Bell, C. Kurschner, C. Martin, M. Smoot, S. Sahasrabudhe, D.L. Barber, S.K. Chanda, T. Ideker, A human MAP kinase interactome, Nat. Methods 7 (10) (2010) 801–805, http://dx.doi.org/10.1038/nmeth.1506.