3-25-2013

Fuzzy complex formation between the intrinsically disordered prothymosin α and the Kelch domain of Keap1 involved in the oxidative stress response.

Halema Khan

Elio A Cino

Anne Brickenden

Jingsong Fan

Daiwen Yang

See next page for additional authors

Follow this and additional works at: <https://ir.lib.uwo.ca/biochempub>

Part of the [Biochemistry Commons](#), and the [Structural Biology Commons](#)

Citation of this paper:

Khan, Halema; Cino, Elio A; Brickenden, Anne; Fan, Jingsong; Yang, Daiwen; and Choy, Wing-Yiu, "Fuzzy complex formation between the intrinsically disordered prothymosin α and the Kelch domain of Keap1 involved in the oxidative stress response." (2013). *Biochemistry Publications*. 156.
<https://ir.lib.uwo.ca/biochempub/156>

Authors

Halema Khan, Elio A Cino, Anne Brickenden, Jingsong Fan, Daiwen Yang, and Wing-Yiu Choy

Fuzzy Complex Formation between the Intrinsically Disordered Prothymosin α and the Kelch Domain of Keap1 Involved in the Oxidative Stress Response

**Halema Khan[†], Elio A. Cino[†], Anne Brickenden[†], Jingsong Fan[‡], Daiwen Yang[‡], and
Wing-Yiu Choy^{†*}**

*[†]Department of Biochemistry, The University of Western Ontario, London, Ontario,
Canada N6A 5C1. [‡]Department of Biological Sciences, National University of Singapore,
14 Science Drive 4, Singapore 117543*

*Corresponding author

E-mail: jchoy4@uwo.ca

Tel: 1-519-661-3161

Fax: 1-519-661-3175

Link to the published article in Journal of Molecular Biology:

[Khan et al. \(2013\) *J Mol Biol*, 425, 1011-1027](#)

(<http://www.sciencedirect.com/science/article/pii/S0022283613000089>)

Abstract

Kelch-like ECH-associated protein 1 (Keap1) is an inhibitor of nuclear factor erythroid 2-related factor 2 (Nrf2), a key transcription factor for cytoprotective gene activation in the oxidative stress response. Under unstressed conditions, Keap1 interacts with Nrf2 in the cytoplasm via its Kelch domain and suppresses the transcriptional activity of Nrf2. During oxidative stress, Nrf2 is released from Keap1 and is translocated into the nucleus, where it interacts with the small Maf protein to initiate gene transcription. Prothymosin alpha (ProT α), an intrinsically disordered protein, also interacts with the Kelch domain of Keap1 and mediates the import of Keap1 into the nucleus to inhibit Nrf2 activity. To gain a molecular basis understanding of the oxidative stress response mechanism, we have characterized the interaction between ProT α and the Kelch domain of Keap1 by using nuclear magnetic resonance spectroscopy (NMR), isothermal titration calorimetry (ITC), peptide array analysis, site-directed mutagenesis, and molecular dynamic (MD) simulations. The results of NMR chemical shift mapping, amide hydrogen exchange, and spin relaxation measurements revealed that ProT α retains a high level of flexibility, even in the bound state with Kelch. This finding is in agreement with the observations from the MD simulations of the ProT α -Kelch complex. Mutational analysis of ProT α , guided by peptide array data and ITC, further pinpointed that the region ³⁸NANEENGE⁴⁵ of ProT α is crucial for the interaction with the Kelch domain, while the flanking residues play relatively minor roles in the affinity of binding.

Keywords: nuclear magnetic resonance; protein dynamics; protein-protein interaction; isothermal titration calorimetry; molecular dynamic simulation

Abbreviations used: CD, circular dichroism; ESI-MS, electrospray ionization mass spectrometry; HSQC, heteronuclear single quantum coherence; IDP, intrinsically disordered protein; ITC, isothermal titration calorimetry; Keap1, Kelch-like ECH-associated protein 1; MD, molecular dynamic; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; Nrf2, Nuclear factor erythroid 2-related factor 2; ProT α , Prothymosin α ; R_1 , longitudinal relaxation rate; $R_{1\rho}$, relaxation rates in rotating frame; R_2 , longitudinal relaxation rate

Introduction

Prothymosin α (ProT α) is a 12-kDa acidic protein with multiple biological functions¹⁻³. It is highly conserved in mammals, is widely distributed in various tissues, and has a distinct amino acid composition¹⁻³. Its protein sequence lacks aromatic and Cys residues, and about 50% of the sequence is composed of either Glu or Asp. These acidic residues are mostly localized in the center of the protein sequence¹⁻⁵. A nuclear localization signal, TKKQK, exists in the C-terminal region^{6,7}. Early studies of ProT α had shown that the protein is unstructured under non-denaturing conditions but can adopt a partially folded conformation at acidic pH^{8,9}. We have extensively characterized the structural and dynamic properties of this protein under physiological buffer conditions by using nuclear magnetic resonance (NMR) spectroscopy and electrospray ionization mass spectrometry (ESI-MS)¹⁰. The results clearly demonstrated that ProT α is largely disordered and extremely dynamic, with only a slight propensity for β -strand structure. However, the protein can adopt more compact conformations upon binding to zinc ions^{10,11}.

ProT α is an oncoprotein expressed at high levels in proliferating cells. It increases cell proliferation by shortening the G1 phase of the cell cycle¹². Studies have shown that in the absence of ProT α , cells undergo apoptosis or are unable to divide¹³⁻¹⁵. ProT α also binds histones^{4,16-18}, plays a role in chromatin remodeling^{16,19}, acts as a transcription activator through CREB binding protein^{20,21}, prevents formation of the apoptosome²²⁻²⁵, and facilitates movement of charged particles within the nucleus and surrounding environment²⁶. Further, the protein suppresses HIV-1 by inducing production of type 1 interferon²⁷⁻²⁹, plays a neuroprotective role after ischemic events^{30,31}, and functions as a tumor-associated protein in human colon cancer, lung cancer, urinary tract cancer, and

breast cancer³²⁻³⁶. Recent studies have shown that ProTα also has an important role in the oxidative stress response^{37,38}.

Cells are constantly attacked by reactive oxygen species, electrophiles, carcinogens, and xenobiotics^{39,40}. Pathological conditions, such as neurodegenerative diseases, arthritis, cancer, and cardiovascular diseases, can arise if cells fail to counteract oxidative stress by inducing cytoprotective gene expression³⁹⁻⁴³. Nuclear factor erythroid 2-related factor 2 (Nrf2) is one of the key regulators of the cellular responses to oxidative stress^{40,44,45}. Nrf2 is a bZIP transcription activator of cytoprotective genes^{40,44}. The protein is composed of six domains, Neh1 to Neh6⁴⁶. The N-terminal Neh2 domain, which is intrinsically disordered⁴⁷, is involved in maintaining homeostatic levels of Nrf2 through binding to Keap1 (Kelch-like ECH-associated protein 1)⁴⁶. Keap1, an adaptor protein in the Cul3-based E3 ubiquitin ligase complex⁴⁸, is composed of five distinct regions as follows: i) N-terminal region, ii) BTB (Broad complex, Tramtrack, and Bric-a-Brac), iii) IVR (intervening region), iv) DGR (double glycine repeat), and v) C-terminal region^{46,48}. The DGR and C-terminal regions of Keap1, known as the Kelch domain, form a six-bladed β-propeller structure, which harbors the binding site of the Neh2 domain of Nrf2⁴⁹.

Under basal cellular conditions, the Neh2 domain of Nrf2 interacts with the Kelch domains of the Keap1 dimer via the ETGE and DLG motifs^{47,50}. This interaction results in ubiquitination and ultimately, the degradation of Nrf2⁴⁸. When the cell is under oxidative stress, Keap1 is no longer able to bind Nrf2 and the latter is translocated to the nucleus, where it interacts with the small Maf protein to activate transcription of cytoprotective genes⁴⁰.

ProTα plays a crucial role in the Nrf2 signaling pathway by mediating the nuclear import of Keap1/Cul3-Rbx1 in order to lower Nrf2 levels in the nucleus, allowing the cell to return to normal conditions^{38,51}. Once in the nucleus, ProTα dissociates from the Keap1/Cul3-Rbx1 complex, allowing Nrf2 to bind to the complex for degradation^{38,51}.

The interaction between mouse ProTα and Keap1 has been studied previously. Using the yeast two hybrid system and different protein constructs, amino acids 32 to 52 in the mouse ProTα have been identified to be crucial for the interaction with the Kelch domain of Keap1³⁷. Padmanabhan *et al.* expanded on these findings by performing structural analysis on the mouse ProTα-Kelch complex⁵². A 16-mer peptide composed of ³⁹AQNEENGQEADNEVD⁵⁴ of mouse ProTα was crystallized with the Kelch domain⁵². The structure reveals that the negatively charged ProTα peptide binds to the basic bottom face of the β-propeller structure of Kelch, forming many electrostatic interactions, especially via the ENGE motif of ProTα. It is noteworthy that the electron density corresponding to residues 49-54 of the ProTα peptide was missing, suggesting this part of the peptide may not adopt a stable conformation upon binding to Kelch⁵². Compared to the structures of Neh2-Kelch⁵²⁻⁵⁴, the EENG motif of ProTα and the EETGE motif in the Neh2 domain bind to the same site on the Kelch domain⁵².

Although ProTα has multiple biological functions and many targets, very limited structural information about this disordered protein is available, especially concerning its target interactions. The crystallographic study of ProTα-Kelch peptide-protein complex is the only detailed structural characterization of ProTα in a bound-state form⁵². However, the molecular mechanism by which full-length ProTα binds to Kelch remains unclear. Many disordered proteins undergo disorder-to-order transition upon binding to their

targets⁵⁵⁻⁵⁷. These protein-protein interactions are frequently enthalpy driven because the loss of conformational freedom of an intrinsically disordered protein (IDP) upon target binding is usually significant, with a resulting unfavorable entropy change⁵⁷⁻⁶¹. IDPs can also interact with targets via preformed structural elements, perhaps lessening the entropy loss and modulating the affinity of binding⁶²⁻⁶⁶. Intriguingly, some IDPs remain disordered even in complexes with targets⁶⁷⁻⁶⁹. Due to the crucial roles ProTα and Keap1 have in the oxidative stress response, in this study, we used experimental and computational approaches to delineate the molecular basis of their binding. The result will not only lead to a better understanding how these two proteins function in the Nrf2 pathway, it will also provide insight into designing drugs that specifically target the interactions between Keap1 and its disordered partners⁷⁰⁻⁷⁴.

Results

Mapping the Kelch-binding region on ProT α by NMR spectroscopy

The Keap1-binding region on the full-length human ProT α was identified by using NMR spectroscopy. A two-dimensional ^1H - ^{15}N heteronuclear single quantum coherence (HSQC) spectrum of ^{15}N -labeled ProT α was collected (Fig. 1a). The lack of chemical shift dispersion in the ^1H dimension of the spectrum indicates that the protein was intrinsically disordered in the absence of Kelch. This finding agrees qualitatively with the disorder predictions of ProT α by PONDR-FIT⁷⁵ and metaPRDOS⁷⁶ (Fig. S1), suggesting that the protein is largely unstructured. These results are also consistent with our previous structural characterization of ProT α carried out in a buffer with lower ionic strength^{10,77}. Potential binding regions of disordered ProT α were also predicted by using ANCHOR⁷⁸, and MoRFpred⁷⁹. Based on the protein sequence, ANCHOR predicted that the binding sites of ProT α are scattered throughout a large part of the sequence (residues 1-57 and 76-99), whereas MoRFpred anticipated that the molecular recognition features of ProT α are specifically located between residues 38-61 (Fig. S1).

To precisely determine the Kelch-binding region on ProT α , changes in the chemical environment of ProT α upon binding to the Kelch domain were then probed by a series of ^1H - ^{15}N HSQC experiments acquired with increasing concentrations of unlabeled Kelch (Fig. 1a). Additions of Kelch led to the perturbations of several ProT α backbone $^1\text{H}/^{15}\text{N}$ resonances. Significant chemical shift changes or attenuations in peak intensity were observed for N36, G37, N38, A39, N40, N43, G44, D49, and N50, as well as several unassigned residues¹⁰. Notably, the dispersion of NMR signals in the ^1H dimension remained very narrow, indicating that ProT α did not undergo global disorder-

to-order transition upon binding to the Kelch domain of Keap1. Aside from the resonance changes mentioned above, most of the peaks in the ^1H - ^{15}N HSQC spectra remain unchanged in position. This result strongly suggests that the dynamic nature of the rest of ProT α was retained upon target binding.

To identify the conformational changes of ProT α induced by the binding of Kelch, backbone $^{13}\text{C}\alpha/\beta$ chemical shifts were analyzed using the residue-specific secondary structure propensity (SSP) program⁸⁰. Based on the differences between the observed chemical shifts and random coil values, a single score between 1 and -1, which is indicative of stable α -helical structure and β -strand structure, respectively, was assigned to each residue. SSP scores of ProT α in the absence and presence of Kelch were compared. The result (Fig. 1b) shows that moderate changes occurred in SSPs of ProT α in the segment spanning residues 36-52; however, the rest of the protein had similar SSP scores in the absence and presence of Kelch. Notably, in the bound state, amino acids in region 35-53 had SSP scores ranging only from -0.34 to 0.17, indicating a lack of stable secondary structure in this region.

Changes in the chemical environment of Kelch upon binding to ProT α were also monitored by HSQC experiments. In the presence of two molar equivalents of the full-length ProT α , many resonance signals of the Kelch domain display noticeable changes in chemical shift or significant decreases in intensity. The same set of peaks are perturbed in the presence of two molar equivalents of the ProT α peptide spanning residues 33-52, albeit to different extents (Fig. S2). Nevertheless, the results clearly illustrate that the full-length ProT α and peptide bind to the same site on the Kelch domain. Notably, similar spectral changes were observed for the interaction between the mouse Kelch and Neh2

peptide⁷⁴. Resonance signals of the mouse Kelch that showed large chemical shift changes or intensity attenuations upon binding to the Neh2 peptide were consistent with the binding interface determined by X-ray crystallography (PDB 1X2R). More detailed structural and dynamic characterization of the 32-kDa human Kelch domain and its target-binding by NMR is in progress.

Disordered ProTα forms a fuzzy complex with the Kelch domain of Keap1

To further characterize the dynamic properties of the ProTα-Kelch complex, backbone ¹⁵N spin relaxation and steady-state ¹H-¹⁵N NOE experiments were performed on the free and Kelch-bound states of ProTα. These experiments are commonly used in studying protein dynamics and are particularly valuable for probing motions on the ps-ns timescale⁸¹. Fig. 2 shows the ¹⁵N R_1 , R_2 , and steady-state ¹H-¹⁵N NOEs measured for ¹⁵N-labeled ProTα (150 μM) in the absence and presence of 300 μM of unlabeled Kelch. Forty-three assigned residues (out of the 110 ProTα residues) with well resolved peaks were included in the data analysis. In the absence of Kelch, all of these 43 residues have R_1 values within a narrow range of 0.87 to 3.13 s⁻¹ (2.45 s⁻¹ in average) (Fig. 2a). Interestingly, addition of Kelch led to insignificant changes in the R_1 values of ProTα (0.85 – 3.06 s⁻¹, average = 2.47 s⁻¹). In the absence of Kelch, the observed R_2 ranged from 0.5 to 5.3 s⁻¹. Notably, for a globular protein with similar molecular weight of ProTα, residues in structured regions are expected to have $R_2 \sim 10$ s⁻¹⁸². Upon addition of Kelch, small, yet significant, increases in R_2 values were observed for most of the residues (Fig. 2b). Larger increases were noted for residues in the region 30-53, which are involved in

the interaction with Kelch, as identified by the chemical shift mapping result mentioned in the previous section.

The steady-state ^1H - ^{15}N NOE is sensitive to fast internal motion, a good indicator of protein mobility on the ps-ns timescale^{81,83}. Generally, large negative NOE values are observed in flexible regions of a protein⁸³. In the free state, NOE values of ProT α were mostly negative, which is consistent with the disordered nature of this protein (Fig. 2c). In the Kelch-bound state, increases in the NOE values were observed for residues 30 to 59 (-0.49 to 0.17), indicating that this region became more restricted in motion (Fig. 2c). Notably, for residues in a structured complex of size ~46 kDa, the NOE values are expected to be ~0.8. Note that the errors of the measurements are significant for those peaks with NOEs close to zero.

To further investigate the nature of the ProT α -Kelch interaction, we conducted NMR amide hydrogen exchange experiments to probe the changes in solvent accessibility of ProT α upon binding to Kelch⁸⁴. Protection factor, which is defined as the ratio of the intrinsic to the measured exchange rate, was determined for backbone amides of ProT α (with isolated signals) in the absence and presence of Kelch (Fig. 3). The low protection factor values of ProT α (0.3 – 5.5; average of 1.6) in the absence of Kelch indicate that the backbone amides were largely exposed to solvent, as expected for a highly unstructured protein like ProT α . Interestingly, no dramatic increases in protection factor values were observed upon addition of Kelch. Even though moderate increases in protection factor values for N38, E45, N50, and G60, were observed, these values were significantly smaller than those observed in folded proteins. For instance, residues located in the secondary structure elements of lysozyme have amide protection factors of 10^3 - 10^8 ⁸⁵.

Identifying crucial residues of ProT α for Kelch binding

Permutation peptide arrays were used to examine the roles that individual residues in ProT α play in the interaction with the Kelch domain. A series of 20-mer peptides, spanning residues 38-57 of human ProT α , were synthesized on a functionalized cellulose membrane and then probed with the Kelch protein. Each of the wild-type residues 38-57 was substituted with any of the 19 amino acids (Fig. 4). Cys was excluded in the permutations to avoid formation of covalent linkages to Kelch during probing, which would result in false positives. Based on the signal intensities observed in the peptide array experiments, mutations that significantly attenuated the binding of the ProT α peptides to Kelch were identified. Many mutations in the ³⁸NANEENGINE⁴⁵ region were found to be detrimental to Kelch binding, whereas replacing other residues outside this region had negligible effects on protein-peptide interactions. In particular, mutations of residues E42, G44, and E45 resulted in loss of interaction with Kelch, indicating that these three amino acids are critical for the binding. On the other hand, mutation of N43 in the ⁴²ENGINE⁴⁵ motif to a Thr or Ser enhanced the binding to Kelch.

In the peptide array experiment, the E41P mutation also resulted in a higher affinity of binding between Kelch and the ProT α peptide. The side chain of E41 forms hydrogen bonds across the binding interface with Kelch⁵². Replacing the glutamate in this position with a non-polar proline was expected to lead to loss of favorable interactions. Therefore, based on the peptide array result, it was unclear why the E41P mutant of ProT α displayed a higher affinity to Kelch (see sections in below).

Quantitative analysis of ProT α and Kelch interaction

To obtain quantitative information for the ProT α -Kelch interaction, isothermal titration calorimetry (ITC) was used to measure the thermodynamic parameters of binding. Parameters, including the binding stoichiometry (n), dissociation constant (K_d), enthalpy change (ΔH), and entropy change (ΔS) of binding, were obtained for the wild-type and a series of mutational variants of full-length ProT α . The results are summarized in Table 1. To evaluate the contributions of ProT α residues distant from the Kelch binding motif to the affinity of binding, we first compared the affinity of Kelch to the full-length ProT α and an 11-mer peptide spanning residues 38-48⁵². Fig. 5a & b show the integrated isotherms exhibiting a single-phase curve, which were fitted best to a single-site binding model. The binding stoichiometry obtained for both the peptide and full-length protein are close to 1, indicating 1:1 binding between ProT α and Kelch protein. The result is also in agreement with the stoichiometry derived from sedimentation equilibrium experiments (Fig. S3).

The dissociation constant of the full-length ProT α -Kelch complex was determined to be 2.6 ± 0.4 μ M by ITC. Interestingly, the 11-mer ProT α peptide bound slightly tighter ($K_d = 1.79 \pm 0.09$ μ M) to Kelch compared to full-length ProT α . Despite the similar K_d values, the ΔH and ΔS of these two binding processes were very different. The results (Table 1) indicate a larger entropy loss when the ProT α peptide bound to Kelch. The more negative ΔS value was compensated for by the more favorable ΔH , resulting in similar K_d values for both full-length ProT α and its peptide.

The charge complementarity between ProT α and Kelch is crucial for their interaction. The highly acidic Kelch-binding motif of ProT α binds to the positively

charged bottom face of the β -propeller structure of Kelch, forming multiple salt bridges and hydrogen bonds across the interface⁵². As aforementioned, Karapetian *et al.* demonstrated that the E45G/E51G double mutant of ProT α is not capable of binding to Kelch³⁷. To further delineate the roles that E45 and E51 play in the interface, binding affinities of the E45G, E45A, E51A, and E51G full-length ProT α mutants to Kelch were measured. The results show that E45 is critical for the interaction with Kelch, as the mutation of this residue to Ala or Gly completely abolished the binding to Kelch (Table 1 and Fig. 5f). This observation is consistent with the peptide array result and is in agreement with the crystallographic data⁵², which illustrate that the side-chain of E45 forms hydrogen bonds with multiple amino acids on Kelch. Mutations of E51 (Table 1 and Fig. 5i), on the other hand, had little effect on the binding affinity to Kelch, indicating that this acidic residue plays only a minor role in mediating the interaction of these two proteins. Our peptide array result clarifies that E45, but not E51, is indispensable for the protein-protein interaction.

The peptide array result also indicates that the mutation of E47, which is located close to the ⁴²ENGE⁴⁵ motif, to any other amino acid (except for R, D, and W) leads to a decrease in binding affinity to Kelch. Based on the charge complementarity at the ProT α and Kelch interaction sites, the result of the E47R mutation is unexpected. To validate this finding, binding affinities of E47R and E47A to Kelch were determined by ITC (Table 1 and Fig. 5g & h). Results show that while the Ala mutation caused a 55-fold decrease in the binding affinity to Kelch, the increase in K_d for E47R was less significant ($K_d = 30 \mu\text{M}$ compared to $2.6 \mu\text{M}$ for the wild-type).

E41 is another acidic residue that showed an interesting result in the peptide array experiment. The crystal structure of the mouse ProT α -Kelch peptide-protein complex reveals that the side-chain of E41 is involved in hydrogen bonding with Kelch⁵². However, mutation of this acidic amino acid to a non-polar proline residue increased the binding affinity to Kelch (Table 1 & Fig. 5d). ITC measurements confirmed that the E41P mutant has a more favorable ΔG_{bind} compared to the wild-type ProT α (-8.57 kcal/mol vs. -7.61 kcal/mol; Table 1), mainly due to the less negative value of the ΔS component.

Based upon the peptide array result, G44 is another residue that is essential for binding Kelch. To confirm this finding, the binding affinity of the G44A mutant to the Kelch domain was determined. As expected, no interaction of this mutant with Kelch was detected by ITC (Fig. 5e). Nevertheless, this specific mutation was very helpful in validating the backbone assignments around this amino acid. The other two mutants that were constructed for the same purpose were N36A and D49A. Both mutational variants bound to Kelch with moderately lower binding affinities (with K_d of 6.5 and 6.7 μM , respectively), compared to wild-type ProT α (Table 1 & Fig. 5c). We point out that even though N36 was excluded from the crystallographic study of the peptide-protein complex of ProT α and Kelch, and the electron density of D49 was absent in the crystal structure solved⁵², both the NMR chemical shift perturbations and ITC results indicate that these two flanking residues also contribute moderately to the binding of ProT α and Kelch.

MD simulations of ProT α -Kelch complex

Molecular dynamic (MD) simulations were performed on the mouse ProT α -Kelch and Neh2-Kelch complexes as a complementary approach to probe the dynamics of the systems (crystal structure of human ProT α -Kelch complex is unavailable thus far). We had included the Neh2-Kelch complex in this study in order to gain insight into the higher binding affinity Neh2 has for Kelch. Fig. S4a & b show the snapshots of ProT α and Neh2 in the Kelch-bound states at 50 ns time intervals. To probe the dynamics of ProT α and Neh2 in their Kelch-bound states quantitatively, circular variances (*C.V.*) of backbone torsion angles of residues around the binding regions were extracted from the 300-ns trajectories. The value of *C.V.* ranges between 0 and 1. The higher the *C.V.* value of an amino acid, the larger is the variability of its torsion angles in the MD simulations. Fig. 6 shows the *C.V.* of individual amino acids close to the Kelch binding motifs of ProT α and Neh2. The plots illustrate that ProT α was systematically more dynamic in the bound state, especially around N38 and G45 (corresponding to N38 and G44 in human ProT α , respectively). Our previous MD simulations of the free-state mouse ProT α demonstrate that residues N41-N44 (N40-N43 in human) are located in a flexible region, but with a propensity of forming a β -turn structure, which is crucial for the bound-state-like β -hairpin formation^{50,65}. Interestingly, the *C.V.* values of N41-N44 observed here in the bound state are similar to those reported for free state, suggesting that the β -turn structure formed by N41-N44 is unstable even in the presence of Kelch.

E41 (in human ProT α) is located in the second position of the β -turn structure mentioned previously. In order to gain insight into the effects of E41P mutation on the β -turn propensity in the free-state, 0.5 μ s MD simulations of a 16-mer ProT α peptide and the E41P mutant in explicit solvent were performed (Fig. S5). Fig. 7 shows the backbone

RMSDs of the β -turn between the bound-state structure and the MD conformations throughout the simulations. The data indicate that even though both the wild-type and E41P ProT α peptides underwent fast conformational inter-conversion during the MD simulations (Fig. S5), both peptides frequently sampled lower RMSD (i.e. < 1.0 Å) bound-state-like conformations. The results suggest that the E41P mutant may still retain the high tendency of forming a bound-state-like β -hairpin structure around the binding motif. Despite that these MD simulations provide insight into the conformational propensities of the wild-type and E41P ProT α peptides, significantly longer trajectories and more advanced simulation techniques are necessary to further investigate factors that stabilize the bound-state-like conformations and to improve the convergence^{65,66}.

Discussion

To better understand the molecular mechanism underlying the oxidative stress response, in this study, we scrutinized the mechanism by which the intrinsically disordered ProT α interacts with the Kelch domain of Keap1. Many studies have shown that IDPs undergo disorder-to-order transition upon binding to their targets through the coupled binding and folding process^{55,60}. Remarkably, our studies of the ProT α -Kelch complex have revealed that ProT α retains its disordered nature, even in the bound state. This finding is supported by the observations that the resonance signals of ProT α in the ¹H-¹⁵N HSQC experiments were clustered in a narrow range from 7.9 ppm to 8.6 ppm, both in the free and bound states (Fig. 1a). Significant chemical shift changes were observed only for amino acids 35-50. Most of these residues have also been identified as crucial in mediating the interaction between the mouse homologs of ProT α and Kelch^{37,52}. The crystal structure of mouse Kelch in complex with a ProT α peptide spanning amino acids 39-54 (corresponding to residues 38-53 in human) reveals that this particular segment of ProT α forms a β -turn conformation upon binding to the bottom side of the β -propeller structure of Kelch⁵². The binding interface is found to be highly complementary in charge, with the acidic amino acids in ProT α playing crucial roles in interacting with multiple Arg residues located on the binding surface of Kelch⁵². In particular, the EENGINE motif in ProT α (residues 41-45 in human ProT α) was shown to be essential for the Kelch binding⁵². Our NMR titration results are in agreement with these previous findings, and importantly, we show that several other residues flanking the EENGINE motif also display attenuation in resonance signals, indicating they may also be involved in Kelch binding.

Compared to the changes in the ^1H - ^{15}N HSQC of ^{15}N -labeled Neh2 upon binding to Kelch, the spectral changes observed for free and bound ProT α were less dramatic. Tong *et al.* have demonstrated that sharp resonance signals of disordered Neh2 were severely broadened out upon addition of Kelch⁴⁷. In the presence of a two molar equivalents of Kelch, most of the resonance signals of Neh2 disappeared, likely due to the significant increase in molecular tumbling time through the formation of a large molecular weight complex or as the result of intermediate conformational exchange between the free and bound states⁴⁷. In contrast, the sharpness of resonance signals prevailed throughout the NMR titration experiments for the majority of residues in ProT α . Moreover, spin relaxation experiments showed that addition of Kelch led to insignificant changes in the R_1 values of ProT α ($0.85 - 3.06 \text{ s}^{-1}$, average = 2.47 s^{-1}). These observed R_1 values are systematically larger than expected for a folded protein with similar molecular weight of ProT α ^{86,87}. For instance, the D39A mutant of the 128-residue apo-IscU has an averaged R_1 value of 1.64 s^{-1} measured under the same magnetic field strength and temperature (i.e., 600 MHz; 298 K), while the averaged R_1 value of the 105-residue PAH2B domain of mSin3B is 1.45 s^{-1} (at 293 K)^{86,87}. The higher R_1 values for ProT α are the result of large amplitude of internal motion on the ps-ns timescale in the protein. Meanwhile, small, yet significant, increases in R_2 values were observed for most of the residues (Fig. 2b) upon addition of Kelch, particularly in the region 30-53. The increases in R_2 suggest that this particular region became more restricted in motion upon binding to Kelch⁸¹. However, the observed R_2 values are still significantly smaller than expected for a 46 kDa (the sum of molecular weights of ProT α and Kelch) complex⁸². Further, the NOE values of ProT α remained predominantly negative upon binding to

Kelch, indicative of a flexible protein lacking secondary structure (Fig. 2c). NOEs of residues in the region directly involved in forming the binding interface with Kelch became positive (increases up to 0.17), but these positive values were still significantly lower than expected for a 46-kDa complex with well-defined structure. Moreover, the small protection factors determined in the amide exchange experiments are in agreement with the low R_2 and small NOE values observed, suggesting that ProT α forms a dynamic complex with Kelch.

Peptide array and ITC experiments were conducted to elucidate further the roles individual amino acids in the binding region of ProT α play in the ‘fuzzy’ interaction with the Kelch domain. The results are compared to the Neh2-Kelch binding. The binding of Nrf2 to Keap1 has been described as a “hinge and latch” mechanism⁵⁰. The disordered Neh2 domain binds to two Kelch domains of a Keap1 dimer via the N-terminal DLG and the C-terminal ETGE motifs, respectively. It has been proposed that the complex formation facilitates the ubiquitination of a Lys residue located between the DLG and ETGE motifs in sequence, leading to the degradation of Nrf2^{47,50}. Previous studies have demonstrated that the ETGE motif of Nrf2 has a higher binding affinity ($K_{a, \text{ETGE}} = 12.4 \times 10^7 \text{ M}^{-1}$) compared to the DLG ($K_{a, \text{DLG}} = 2 \times 10^6 \text{ M}^{-1}$) motif for the Kelch domain^{47,50}. Our ITC results revealed that the binding affinity of ProT α ($K_a = (3.8 \pm 0.6) \times 10^5 \text{ M}^{-1}$) was also significantly weaker than that of the Neh2 ETGE motif. Tong *et al.* reported that the ΔH and $T\Delta S$ of binding between Kelch and mouse Neh2 [$\Delta 1$ -33] with the low affinity DLG motif deleted were -21.2 and -10.2 kcal/mol, respectively⁴⁷. Large and negative values of ΔS are frequently observed for binding involving IDPs mainly due to the significant reduction in conformational entropy⁵⁸. Interestingly, compared to the

thermodynamic parameters obtained in this study, the entropic penalty was lower for the dynamic ProT α -Kelch complex formation (Table 1). On the other hand, the interaction between ProT α and Kelch was weaker based on the values of the enthalpy changes.

Our peptide array result showed evidently that the ⁴²ENGE⁴⁵ motif is critical for binding, since mutations in this region were disruptive. However, exceptions were seen. Mutation of N43 to a Thr or Ser enhanced the binding to Kelch. Intriguingly, with the ETGE motif, the Neh2 domain of Nrf2 has been shown to have a greater affinity for the Kelch domain⁴⁷. Therefore, conversion of the ENGE to an ETGE motif in ProT α resulting in a higher binding affinity is not unexpected. Based on the crystal structures of ProT α -Kelch and Neh2-Kelch peptide-protein complexes, both the Asn in ProT α and the Thr in Neh2 are not directly involved in the interaction with Kelch. Instead, the side-chains of each of these two amino acids play important roles in stabilizing the β -turn conformation of the binding motif by forming hydrogen bonds with the peptide backbone^{47,52}. This finding explains why the N43S mutant of ProT α also showed Kelch binding affinity that is similar to N43T, as the hydroxyl group of Ser may still participate in the intra-molecular hydrogen bonding^{47,52}. Notably, hydrogen bonds between the N/N and D/T pairs in the ⁴⁰NEEN⁴³ and ⁷⁷DEET⁸⁰ motifs of ProT α and Neh2, respectively, also contribute to stabilization of the β -turn conformation of these proteins in their free states^{65,66}.

Meanwhile, the binding affinity between ProT α and Kelch increased by 5-fold when E41, the residue N-terminal to the ENGE motif, of ProT α was mutated to a Pro residue ($K_a = 19.2 \pm 0.8 \times 10^5 \text{ M}^{-1}$). One possible explanation for the higher affinity of the ProT α E41P mutant is that the conformational entropy of the free state is reduced,

lowering the amount of entropic penalty upon binding. The ⁴⁰NEENGINE⁴⁵ region of ProTα adopts β-turn conformations when bound to the Kelch domain^{52,88}. Pro residues are commonly found in position *i*+1 of β-turns in proteins⁸⁹. The cyclic side chain of Pro leads to drastic restriction in the sampling of backbone dihedral angles. Our MD simulation of the E41P peptide (in its free state) suggested that it has a high tendency to form the bound-state-like β-turn structure. Therefore, it seems plausible that the incorporation of a Pro at this position leads to a less severe conformational entropy loss upon binding, resulting in a higher binding affinity. Indeed, lowered free state conformational entropy is thought to play an important role in the high affinity of the Neh2-Kelch domain interaction^{50,65}. Another possible explanation is that, since E41 is involved in forming salt bridges and hydrogen bonds across the binding interface with Kelch, mutating this residue to a non-polar Pro would result in loss of these interactions. The change would render this region of the protein more flexible, thus lowering the reduction in conformational entropy of the system upon binding.

Intriguingly, p62, a protein involved in regulation of autophagy, was recently found to bind to the Kelch domain of Keap1⁹⁰. Crystal structures show that ProTα, p62, and Neh2 all interact with the same site on the Kelch domain^{47,52,90}. Even though the Kelch-binding motif in p62, ³⁵⁰PSTGE³⁵⁴, differs from the ⁴¹EENGINE⁴⁵ and ⁷⁸EETGE⁸² motifs of ProTα and Neh2, respectively, the Gly and Glu residues at the 4th and 5th positions are conserved. This observation is consistent with our results that these two residues are crucial for the interaction with Kelch^{47,52,90}. Interestingly, the corresponding E41 in ProTα is replaced with a Pro residue in p62⁹⁰. ITC experiments performed by Komatsu *et al.* show that the binding affinity of p62 for Kelch is $K_d = 5.4 \pm 0.3 \times 10^5 \text{ M}^{-1}$

¹,⁹⁰ which is slightly higher than what was determined for ProT α and Kelch ($K_a = 3.8 \pm 0.6 \times 10^5 \text{ M}^{-1}$).

E47 is another residue in ProT α that may help to stabilize the bound-state β -hairpin structure. Our ITC results (Table 1) revealed that mutation of this residue to an Ala or Arg decreased the binding affinity to Kelch by 55 fold, and ~ 10 fold, respectively. Since the side-chain of the corresponding Glu was invisible in the crystal structure of mouse ProT α -Kelch peptide-protein complex⁵², it is unclear whether this residue was involved in any intra- or intermolecular interactions. We speculate that the side-chain of E47 may be involved in intra-molecular hydrogen bonding with N38, and the interaction can still be retained in the E47R mutant. The transient interaction between these two residues may also help to stabilize the bound-state-like β -hairpin structure in the free state of ProT α .

We also conclude from our NMR and ITC experiments that even though peptides can be used to mimic specific regions of a protein when studying binding, the target recognition mechanisms of peptides can differ from that of the full-length protein. As seen in Table 1, although the ProT α peptide and protein had similar binding affinities to Kelch, their ΔH and ΔS values differed. Based on the observed chemical shift changes of ProT α upon target binding and the spin relaxation parameters determined in both the free and bound-states, only residues around the binding motif were involved in the interaction with Kelch while the rest of the protein remained highly disordered. One possible explanation for the significant differences in the ΔH and ΔS values is that a greater loss in the conformational entropy of the peptide occurred upon binding to Kelch. This idea is in agreement with our earlier computational study of the free-state structure of full-length

ProT α and peptide using MD simulations⁶⁵. In that work, we had demonstrated that the full-length ProT α has a higher propensity than the peptide in forming the bound-state-like β -hairpin structure, which can result in a less entropic penalty upon coupled binding and folding for the full-length protein. On the other hand, the enthalpy change of the ProT α peptide is more negative compared to the full-length protein. Since the measured ΔH is the sum of the enthalpy changes due to the binding and folding processes, the larger ΔH value of the peptide is likely due to the folding from the less bound-state-like conformations.

Despite the highly dynamic nature of disordered proteins, it has been generally accepted that they will adopt a well-defined structure once bound to the target, at least in the regions that are directly involved in the interaction. Several recent examples, however, challenge this view and suggest that disordered proteins can still be flexible even in the bound states⁶⁷⁻⁶⁹. For instance, the disordered cyclin-dependent kinase inhibitor Sic1 does not undergo a disorder-to-order transition upon binding to Cdc4⁹¹. Instead, NMR studies showed that the multiple and dispersed phosphorylated motifs on Sic1 interact with a single binding site on Cdc4 in a dynamic equilibrium manner. This unusual binding mechanism has been proposed to lead to the ultrasensitivity to phosphorylation, which is crucial for the function of Sic1 in cell cycle regulation⁹¹. Another example of a fuzzy complex is the dynamic interaction between clathrin assembly protein AP180 and clathrin⁶⁹. NMR studies demonstrated that a large fragment of the clathrin binding domain (CBD) of AP180 is disordered. Interestingly, this protein fragment remains unstructured even in the bound form with clathrin. This dynamic binding mode was speculated to be important for the assembly of the clathrin lattice⁶⁹.

In summary, our results here show that only a small region of the highly disordered ProT α is involved in the interaction with the Kelch domain, and the protein does not undergo large conformational change upon binding. These properties may allow ProT α to also interact with other targets simultaneously. In addition, the fuzzy complex formed between the two proteins is probably what allows ProT α to quickly dissociate from Keap1 after the latter is translocated into the nucleus³⁸.

Materials and Methods

Protein Expression and Purification

The expression and purification of human ProT α was reported previously⁷⁷. The pET15b plasmid carrying the human Kelch cDNA, a kind gift from Dr. Mark Hannink at the University of Missouri-Columbia, was transformed into *E. coli* BL21 (DE3) cells (Novagen). The protein was overexpressed in M9 minimal media and induced by adding 0.5 mM isopropyl- β -D-thiogalactopyranoside (IPTG; Bioshop). After 18 h induction at 16 °C, the bacteria were harvested and stored at -20 °C. The His-tagged protein was purified by affinity chromatography using Ni-Sepharose 6 FF beads (Amersham Biosciences). The eluted protein was dialyzed against TE buffer (20 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, 100 mM NaCl). The His-tag was cleaved with human alpha-thrombin (Haematologic Technologies), and protein was further purified with HiLoad Superdex 75 column (GE Healthcare). The final protein sample was dialyzed in 50 mM sodium phosphate buffer, 100 mM NaCl, 1 mM DTT, pH 7.

Peptide synthesis

The ProT α peptide (98% in purity), ³⁸NANEENGQEQA⁴⁸ was ordered from GenScript USA Inc. The lyophilized peptide was dissolved in 50 mM Sodium phosphate buffer, 100 mM NaCl, 1 mM DTT for ITC and NMR experiments.

Site-directed mutagenesis of wild-type ProT α

Wild-type ProT α in pET15b plasmid was subjected to point mutagenesis using the Stratagene Quikchange II site-directed mutagenesis kit. The primers used for each

point mutation are included in Supplemental data. Over expression and purification of mutant variants of ProT α were carried out with the similar procedure used for the wild-type ProT α ⁷⁷.

NMR spectroscopy

All NMR experiments (except the amide hydrogen exchange experiments) were performed on a Varian Inova 600MHz spectrometer equipped with xyz-gradient triple resonance probe at 25 °C in buffer containing 50 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, and 10% (v/v) D₂O (Cambridge Isotope Laboratories). 1 mM of sodium 2, 2-dimethyl-2-silapentane-5-sulfonate (DSS, Sigma) was added to each NMR sample as an internal standard for the chemical shift referencing⁹². All NMR data were analyzed with NMRPipe⁹³, and the spectra were analyzed with NMRView⁹⁴.

For the NMR titration experiments, unlabeled Kelch was titrated into 600 μ L of 150 μ M ¹⁵N-labeled ProT α until a 1:2 molar ratio (ProT α : Kelch) point was reached. ¹H-¹⁵N HSQC spectra were collected for each titration point.

The TROSY-HSQC spectra of Kelch was collected with 200 μ M ¹⁵N-labeled sample. Unlabelled full-length ProT α and ProT α peptide (residues 33-52) were added to achieve a final concentration of 1:2 molar ratio of Kelch to ProT α .

Backbone ¹⁵N longitudinal relaxation rates (R_1), relaxation rates in rotating frame ($R_{1\rho}$), and steady-state ¹H-¹⁵N NOE relaxation parameters of 150 μ M ¹⁵N-labeled ProT α were measured in both the free state and the Kelch bound state (unlabeled 300 μ M Kelch) using identical parameters as described previously¹⁰.

Amide hydrogen exchange (HX) rate measurements were performed at 25 °C on a Bruker Avance 800 MHz spectrometer equipped with cryogenic probe (National University of Singapore) using the method recently proposed by Fan *et al.*⁸⁴. The HX rates of ¹⁵N-labeled ProTα (180 μM) were measured in both the free state and the Kelch-bound state (unlabeled 450 μM Kelch) using the following mixing times: 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 190, 220, 260, and 300 ms. For each mixing time, data sets with and without the suppression of radiation damping were collected in interleaved manner and each FID was acquired with 6 scans. A reference spectrum was collected for each sample using a long interscan delay (10 s) and by removing the first ¹H pulse and the mixing period in the pulse sequence as mentioned in the reference⁸⁴. The data were processed using NMRPipe⁹³. For each isolated peak, signal intensities were fitted to Equation 1 (Equation 9 in the reference⁸⁴) in order to determine the exchange rates.

$$\frac{I_{EX}(t)}{I_{ref}} = \frac{f * k_{ex}}{R_{1H} + k_{ex} - R_{1w}} [\exp(-R_{1w}(t - t_0)) - \exp(-(R_{1H} + k_{ex})(t - t_0))] \quad (1)$$

where f is the fractional steady-state water magnetization; R_{1H} is the relaxation rate of an amide proton; I_{ref} is the equilibrium magnetization of the amide proton, which is equal to two times of the peak intensity of the amide in the reference spectrum when it is recorded with the same number of scans as the other spectra; $I_{EX}(t)$ is the peak intensity of the amide correlation at mixing time t , which is obtained from the difference spectrum of two interleaved data sets; t_0 is the latency interval; and R_{1w} is the longitudinal relaxation rate of water. In the data fitting, f , t_0 , and R_{1w} were set to 0.90, 17.4 ms, and 0.3 s^{-1} , respectively. The protection factor (P) of the backbone amide proton of each amino acid was calculated as the ratio of the intrinsic HX rate (k_{int}) of random coil to the

experimentally determined HX rate. The k_{int} values were estimated with SPHERE web server (<http://www.fccc.edu/research/labs/roder/sphere/sphere.html>)^{95,96}.

Isothermal titration calorimetry (ITC)

ITC experiments were carried out on a MicroCalTM VP-ITC. All purified protein samples were dialyzed into 50 mM sodium phosphate, 100 mM NaCl, 1 mM DTT at pH 7.0. For titrations, purified Kelch and ProT α were prepared to 60 μ M and 790 μ M, respectively, in dialysis buffer, and each was degassed prior to titration. Kelch was loaded into the 1.42 mL cell, and ProT α was loaded into the syringe (300 μ L). Titrations were performed at 25 °C starting with initial injection of 3 μ L, followed by 59 larger injections of 5 μ L, with spacing of 360 s (except for the ProT α peptide with spacing of 120 s). The sample cell was stirred at 300 rpm throughout the experiment. The buffer blank was performed under the same conditions and showed negligible heats of binding. The dissociation constant (K_d), stoichiometry of binding (N), binding enthalpy (ΔH), and entropy (ΔS) were obtained by non-linear least squares fitting of the data to a single-site binding model provided with the data analysis software (Origin 7). Baselines were subtracted from final data using Origin software. All protein concentrations were determined using Lowry (for ProT α) and Bradford (for Kelch) assays. Protein concentrations were confirmed by amino acid analysis (Advanced Protein Technology Centre, The Hospital for Sick Children, Toronto, ON).

Peptide array

A twenty-residue, $^{38}\text{NANEENGQEADNEVDEEEE}^{57}$, ProT α peptide spot array was ordered from the Peptide and Peptide Array Synthesis Facility at the University of Western Ontario. The peptide membrane was hydrated, blocked with 5% skim milk, and then probed with the Kelch domain of Keap1. His Probe-HRP was used to detect the ProT α peptide-Kelch interaction. The membrane was developed with the ECL Lightening Plus kit (Perkin Elmer, Mississauga, ON) before exposing to film.

MD simulations

Several MD simulations were performed in order to gain further insight into the bound and free state dynamics of ProT α and the Neh2 domain of Nrf2. These included simulations of ProT α and the Neh2 domain in complex with the Kelch domain of Keap1 and free state simulations of a 16-mer ProT α peptide and the E41P mutant. All MD simulations were performed using GROMACS version 4⁹⁷ and the GROMOS96 53a6^{98,99} force field.

The starting structures for the bound state simulations were based upon the crystal structures of mouse ProT α and Neh2 domain peptides bound to the mouse Kelch domain (PDB ids: 2Z32 and 1X2R, respectively)^{52,54}. The rest of the full-length mouse ProT α and Neh2 domain (residues 1-100) were modeled onto these crystal structures. Detailed description of the generation of bound state complexes and the parameters used for the MD simulations are included in Supplemental Data. The simulations were run for 300 ns each.

The starting structures for the free state simulations of the 16-mer ProT α peptide and the E41P mutant were generated using the Crystallography & NMR System (CNS)¹⁰⁰

to generate an extended structure based solely upon the amino acid sequence of a 16-mer human isoform 2 ProT α peptide, ³⁵ANGNANEENGQEADN⁵⁰, and its E41P mutational variant, ³⁵ANGNANPENGEQEADN⁵⁰. Annealing simulations were then performed to relax the fully extended conformation. Unstructured conformations of each peptide from the annealing simulations were used as the starting structures for the MD simulations, which were performed using similar parameters as the bound state simulations. The peptides were centered in boxes large enough so that the minimum distance to the box edge was > 1.0 nm. The systems contained 7620 water molecules, 19 Na⁺ and 14 Cl⁻ ; 8513 waters, 20 Na⁺ and 16 Cl⁻ for the wild-type and E41P mutational variant peptide, respectively. The simulations were run for 500 ns each.

Residue specific dynamics of the ProT α and Neh2 in bound states were analyzed based on the circular variance (*C.V.*) of the ϕ and ψ dihedral angles over time. The *C.V.* is defined as¹⁰¹

$$C.V. = 1 - R/m \quad (2)$$

where *m* is the number of structures included in the analysis, and *R* is calculated using the following equation:

$$R = \sqrt{\frac{1}{2} \left[\left(\sum_{i=1}^m \cos \phi \right)^2 + \left(\sum_{i=1}^m \sin \phi \right)^2 + \left(\sum_{i=1}^m \cos \psi \right)^2 + \left(\sum_{i=1}^m \sin \psi \right)^2 \right]} \quad (3)$$

The value of *C.V.* ranges between 0 and 1. The higher the *C.V.* values, the larger are the ϕ and ψ variability.

Acknowledgements

This work was supported by an Operating Grant (MOP no. 74679) from the Canadian Institutes of Health Research (CIHR). We would like to thank the Biomolecular NMR Facility and the Biomolecular Interaction and Conformation Facility for their assistance and use of the facilities. We also gratefully acknowledge the support of the NMR facility at the National University of Singapore. Computational resources were provided by the Shared Hierarchical Academic Research Computing Network (SHARCNET).

References

1. Haritos, A. A., Goodall, G. J. & Horecker, B. L. (1984). Prothymosin alpha: isolation and properties of the major immunoreactive form of thymosin alpha 1 in rat thymus. *Proc Natl Acad Sci U S A* **81**, 1008-1011.
2. Frillingos, S., Frangou-Lazaridis, M., Seferiadis, K., Hulmes, J. D., Pan, Y. C. & Tsolas, O. (1991). Isolation and partial sequence of goat spleen prothymosin alpha. *Mol Cell Biochem* **108**, 85-94.
3. Pan, L. X., Haritos, A. A., Wideman, J., Komiyama, T., Chang, M., Stein, S., Salvin, S. B. & Horecker, B. L. (1986). Human prothymosin alpha: amino acid sequence and immunologic properties. *Arch Biochem Biophys* **250**, 197-201.
4. Diaz-Jullien, C., Perez-Estevez, A., Covelo, G. & Freire, M. (1996). Prothymosin alpha binds histones in vitro and shows activity in nucleosome assembly assay. *Biochim Biophys Acta* **1296**, 219-227.
5. Segade, F. & Gomez-Marquez, J. (1999). Prothymosin alpha. *Int J Biochem Cell Biol* **31**, 1243-1248.
6. Trumbore, M. W., Wang, R. H., Enkemann, S. A. & Berger, S. L. (1997). Prothymosin alpha in vivo contains phosphorylated glutamic acid residues. *J Biol Chem* **272**, 26394-26404.
7. Rubtsov, Y. P., Zolotukhin, A. S., Vorobjev, I. A., Chichkova, N. V., Pavlov, N. A., Karger, E. M., Evstafieva, A. G., Felber, B. K. & Vartapetian, A. B. (1997). Mutational analysis of human prothymosin alpha reveals a bipartite nuclear localization signal. *FEBS Lett* **413**, 135-141.
8. Gast, K., Damaschun, H., Eckert, K., Schulze-Forster, K., Maurer, H. R., Muller-Frohne, M., Zirwer, D., Czarnecki, J. & Damaschun, G. (1995). Prothymosin alpha: a biologically active protein with random coil conformation. *Biochemistry* **34**, 13211-13218.
9. Uversky, V. N., Gillespie, J. R., Millett, I. S., Khodyakova, A. V., Vasiliev, A. M., Chernovskaya, T. V., Vasilenko, R. N., Kozlovskaya, G. D., Dolgikh, D. A., Fink, A. L., Doniach, S. & Abramov, V. M. (1999). Natively unfolded human prothymosin alpha adopts partially folded collapsed conformation at acidic pH. *Biochemistry* **38**, 15009-15016.
10. Yi, S., Boys, B. L., Brickenden, A., Konermann, L. & Choy, W. Y. (2007). Effects of zinc binding on the structure and dynamics of the intrinsically disordered protein prothymosin alpha: evidence for metalation as an entropic switch. *Biochemistry* **46**, 13120-13130.
11. Uversky, V. N., Gillespie, J. R., Millett, I. S., Khodyakova, A. V., Vasilenko, R. N., Vasiliev, A. M., Rodionov, I. L., Kozlovskaya, G. D., Dolgikh, D. A., Fink, A. L., Doniach, S., Permyakov, E. A. & Abramov, V. M. (2000). Zn(2+)-mediated structure formation and compaction of the "natively unfolded" human prothymosin alpha. *Biochem Biophys Res Commun* **267**, 663-668.
12. Rodriguez, P., Vinuela, J. E., Alvarez-Fernandez, L., Buceta, M., Vidal, A., Dominguez, F. & Gomez-Marquez, J. (1998). Overexpression of prothymosin alpha accelerates proliferation and retards differentiation in HL-60 cells. *Biochem J* **331 (Pt 3)**, 753-761.

13. Sburlati, A. R., Manrow, R. E. & Berger, S. L. (1991). Prothymosin alpha antisense oligomers inhibit myeloma cell division. *Proc Natl Acad Sci U S A* **88**, 253-257.
14. Rodriguez, P., Vinuela, J. E., Alvarez-Fernandez, L. & Gomez-Marquez, J. (1999). Prothymosin alpha antisense oligonucleotides induce apoptosis in HL-60 cells. *Cell Death Differ* **6**, 3-5.
15. Evstafieva, A. G., Belov, G. A., Kalkum, M., Chichkova, N. V., Bogdanov, A. A., Agol, V. I. & Vartapetian, A. B. (2000). Prothymosin alpha fragmentation in apoptosis. *FEBS Lett* **467**, 150-154.
16. Karetso, Z., Sandaltzopoulos, R., Frangou-Lazaridis, M., Lai, C. Y., Tsolas, O., Becker, P. B. & Papamarcaki, T. (1998). Prothymosin alpha modulates the interaction of histone H1 with chromatin. *Nucleic Acids Res* **26**, 3111-3118.
17. Papamarcaki, T. & Tsolas, O. (1994). Prothymosin alpha binds to histone H1 in vitro. *FEBS Lett* **345**, 71-75.
18. Martic, G., Karetso, Z., Kefala, K., Politou, A. S., Clapier, C. R., Straub, T. & Papamarcaki, T. (2005). Parathymosin affects the binding of linker histone H1 to nucleosomes and remodels chromatin structure. *J Biol Chem* **280**, 16143-16150.
19. Karetso, Z., Martic, G., Tavoulari, S., Christoforidis, S., Wilm, M., Gruss, C. & Papamarcaki, T. (2004). Prothymosin alpha associates with the oncoprotein SET and is involved in chromatin decondensation. *FEBS Lett* **577**, 496-500.
20. Karetso, Z., Kretsovali, A., Murphy, C., Tsolas, O. & Papamarcaki, T. (2002). Prothymosin alpha interacts with the CREB-binding protein and potentiates transcription. *EMBO Rep* **3**, 361-366.
21. Subramanian, C., Hasan, S., Rowe, M., Hottiger, M., Orre, R. & Robertson, E. S. (2002). Epstein-Barr virus nuclear antigen 3C and prothymosin alpha interact with the p300 transcriptional coactivator at the CH1 and CH3/HAT domains and cooperate in regulation of transcription and histone acetylation. *J Virol* **76**, 4699-4708.
22. Jiang, X., Kim, H. E., Shu, H., Zhao, Y., Zhang, H., Kofron, J., Donnelly, J., Burns, D., Ng, S. C., Rosenberg, S. & Wang, X. (2003). Distinctive roles of PHAP proteins and prothymosin-alpha in a death regulatory pathway. *Science* **299**, 223-226.
23. Qi, X., Wang, L. & Du, F. (2010). Novel small molecules relieve prothymosin alpha-mediated inhibition of apoptosome formation by blocking its interaction with Apaf-1. *Biochemistry* **49**, 1923-1930.
24. Piacentini, M., Evangelisti, C., Mastroberardino, P. G., Nardacci, R. & Kroemer, G. (2003). Does prothymosin-alpha act as molecular switch between apoptosis and autophagy? *Cell Death Differ* **10**, 937-939.
25. Malicet, C., Giroux, V., Vasseur, S., Dagorn, J. C., Neira, J. L. & Iovanna, J. L. (2006). Regulation of apoptosis by the p8/prothymosin alpha complex. *Proc Natl Acad Sci U S A* **103**, 2671-2676.
26. Enkemann, S. A., Ward, R. D. & Berger, S. L. (2000). Mobility within the nucleus and neighboring cytosol is a key feature of prothymosin-alpha. *J Histochem Cytochem* **48**, 1341-1355.

27. Kubota, S., Adachi, Y., Copeland, T. D. & Oroszlan, S. (1995). Binding of human prothymosin alpha to the leucine-motif/activation domains of HTLV-I Rex and HIV-1 Rev. *Eur J Biochem* **233**, 48-54.
28. Mosoian, A., Teixeira, A., Burns, C. S., Khitrov, G., Zhang, W., Gusella, L., Klotman, P. & Klotman, M. (2007). Influence of prothymosin-alpha on HIV-1 target cells. *Ann N Y Acad Sci* **1112**, 269-285.
29. Mosoian, A., Teixeira, A., Burns, C. S., Sander, L. E., Gusella, G. L., He, C., Blander, J. M., Klotman, P. & Klotman, M. E. (2010). Prothymosin-alpha inhibits HIV-1 via Toll-like receptor 4-mediated type I interferon induction. *Proc Natl Acad Sci U S A* **107**, 10178-10183.
30. Ueda, H., Fujita, R., Yoshida, A., Matsunaga, H. & Ueda, M. (2007). Identification of prothymosin-alpha1, the necrosis-apoptosis switch molecule in cortical neuronal cultures. *J Cell Biol* **176**, 853-862.
31. Fujita, R., Ueda, M., Fujiwara, K. & Ueda, H. (2009). Prothymosin-alpha plays a defensive role in retinal ischemia through necrosis and apoptosis inhibition. *Cell Death Differ* **16**, 349-358.
32. Gou, L. T., Tong, A. P., Yan, F., Yuan, Z., He, F., Wang, W., Zhou, Y., Chen, L. J., Tang, M. H. & Yang, J. L. (2009). Altered protein-expressing profile in hPNAS4-induced apoptosis in A549 human lung adenocarcinoma cells. *J Cell Biochem* **108**, 1211-1219.
33. Tsai, Y. S., Jou, Y. C., Lee, G. F., Chen, Y. C., Shiau, A. L., Tsai, H. T., Wu, C. L. & Tzai, T. S. (2009). Aberrant prothymosin-alpha expression in human bladder cancer. *Urology* **73**, 188-192.
34. Traub, F., Jost, M., Hess, R., Schorn, K., Menzel, C., Budde, P., Schulz-Knappe, P., Lamping, N., Pich, A., Kreipe, H. & Tammen, H. (2006). Peptidomic analysis of breast cancer reveals a putative surrogate marker for estrogen receptor-negative carcinomas. *Lab Invest* **86**, 246-253.
35. Dominguez, F., Magdalena, C., Cancio, E., Roson, E., Paredes, J., Loidi, L., Zalvide, J., Fraga, M., Forteza, J., Regueiro, B. J. & et al. (1993). Tissue concentrations of prothymosin alpha: a novel proliferation index of primary breast cancer. *Eur J Cancer* **29A**, 893-897.
36. Tsitsiloni, O. E., Stiakakis, J., Koutselinis, A., Gogas, J., Markopoulos, C., Yialouris, P., Bekris, S., Panoussopoulos, D., Kiortsis, V., Voelter, W. & et al. (1993). Expression of alpha-thymosins in human tissues in normal and abnormal growth. *Proc Natl Acad Sci U S A* **90**, 9504-9507.
37. Karapetian, R. N., Evstafieva, A. G., Abaeva, I. S., Chichkova, N. V., Filonov, G. S., Rubtsov, Y. P., Sukhacheva, E. A., Melnikov, S. V., Schneider, U., Wanker, E. E. & Vartapetian, A. B. (2005). Nuclear oncoprotein prothymosin alpha is a partner of Keap1: implications for expression of oxidative stress-protecting genes. *Mol Cell Biol* **25**, 1089-1099.
38. Niture, S. K. & Jaiswal, A. K. (2009). Prothymosin-alpha mediates nuclear import of the INrf2/Cul3 Rbx1 complex to degrade nuclear Nrf2. *J Biol Chem* **284**, 13856-13868.
39. Primiano, T., Sutter, T. R. & Kensler, T. W. (1997). Antioxidant-inducible genes. *Adv Pharmacol* **38**, 293-328.

40. Itoh, K., Chiba, T., Takahashi, S., Ishii, T., Igarashi, K., Katoh, Y., Oyake, T., Hayashi, N., Satoh, K., Hatayama, I., Yamamoto, M. & Nabeshima, Y. (1997). An Nrf2/small Maf heterodimer mediates the induction of phase II detoxifying enzyme genes through antioxidant response elements. *Biochem Biophys Res Commun* **236**, 313-322.
41. Kim, D. & Guengerich, F. P. (2005). Cytochrome P450 activation of arylamines and heterocyclic amines. *Annu Rev Pharmacol Toxicol* **45**, 27-49.
42. Schulze, P. C. & Lee, R. T. (2005). Oxidative stress and atherosclerosis. *Curr Atheroscler Rep* **7**, 242-248.
43. Pratico, D., Rokach, J., Lawson, J. & FitzGerald, G. A. (2004). F2-isoprostanes as indices of lipid peroxidation in inflammatory diseases. *Chem Phys Lipids* **128**, 165-171.
44. Venugopal, R. & Jaiswal, A. K. (1996). Nrf1 and Nrf2 positively and c-Fos and Fra1 negatively regulate the human antioxidant response element-mediated expression of NAD(P)H:quinone oxidoreductase1 gene. *Proc Natl Acad Sci U S A* **93**, 14960-14965.
45. Itoh, K., Tong, K. I. & Yamamoto, M. (2004). Molecular mechanism activating Nrf2-Keap1 pathway in regulation of adaptive response to electrophiles. *Free Radic Biol Med* **36**, 1208-1213.
46. Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., Igarashi, K., Engel, J. D. & Yamamoto, M. (1999). Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. *Genes Dev* **13**, 76-86.
47. Tong, K. I., Katoh, Y., Kusunoki, H., Itoh, K., Tanaka, T. & Yamamoto, M. (2006). Keap1 recruits Neh2 through binding to ETGE and DLG motifs: characterization of the two-site molecular recognition model. *Mol Cell Biol* **26**, 2887-2900.
48. Kobayashi, A., Kang, M. I., Okawa, H., Ohtsuji, M., Zenke, Y., Chiba, T., Igarashi, K. & Yamamoto, M. (2004). Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2. *Mol Cell Biol* **24**, 7130-7139.
49. Li, X., Zhang, D., Hannink, M. & Beamer, L. J. (2004). Crystallization and initial crystallographic analysis of the Kelch domain from human Keap1. *Acta Crystallogr D Biol Crystallogr* **60**, 2346-2348.
50. Tong, K. I., Kobayashi, A., Katsuoka, F. & Yamamoto, M. (2006). Two-site substrate recognition model for the Keap1-Nrf2 system: a hinge and latch mechanism. *Biol Chem* **387**, 1311-1320.
51. Niture, S. K., Kaspar, J. W., Shen, J. & Jaiswal, A. K. (2010). Nrf2 signaling and cell survival. *Toxicol Appl Pharmacol* **244**, 37-42.
52. Padmanabhan, B., Nakamura, Y. & Yokoyama, S. (2008). Structural analysis of the complex of Keap1 with a prothymosin alpha peptide. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**, 233-238.
53. Lo, S. C., Li, X., Henzl, M. T., Beamer, L. J. & Hannink, M. (2006). Structure of the Keap1:Nrf2 interface provides mechanistic insight into Nrf2 signaling. *EMBO J* **25**, 3605-3617.

54. Padmanabhan, B., Tong, K. I., Ohta, T., Nakamura, Y., Scharlock, M., Ohtsuji, M., Kang, M. I., Kobayashi, A., Yokoyama, S. & Yamamoto, M. (2006). Structural basis for defects of Keap1 activity provoked by its point mutations in lung cancer. *Mol Cell* **21**, 689-700.
55. Wright, P. E. & Dyson, H. J. (2009). Linking folding and binding. *Curr Opin Struct Biol* **19**, 31-38.
56. Dyson, H. J. & Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* **6**, 197-208.
57. Wright, P. E. & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* **293**, 321-331.
58. Meszaros, B., Tompa, P., Simon, I. & Dosztanyi, Z. (2007). Molecular principles of the interactions of disordered proteins. *J Mol Biol* **372**, 549-561.
59. Fink, A. L. (2005). Natively unfolded proteins. *Curr Opin Struct Biol* **15**, 35-41.
60. Sugase, K., Dyson, H. J. & Wright, P. E. (2007). Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021-1025.
61. Bordelon, T., Montegudo, S. K., Pakhomova, S., Oldham, M. L. & Newcomer, M. E. (2004). A disorder to order transition accompanies catalysis in retinaldehyde dehydrogenase type II. *J Biol Chem* **279**, 43085-43091.
62. Tsai, C. J., Ma, B., Sham, Y. Y., Kumar, S. & Nussinov, R. (2001). Structured disorder and conformational selection. *Proteins* **44**, 418-427.
63. Fuxreiter, M., Simon, I., Friedrich, P. & Tompa, P. (2004). Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J Mol Biol* **338**, 1015-1026.
64. Kumar, S., Ma, B., Tsai, C. J., Sinha, N. & Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci* **9**, 10-19.
65. Cino, E. A., Wong-ekkabut, J., Karttunen, M. & Choy, W. Y. (2011). Microsecond molecular dynamics simulations of intrinsically disordered proteins involved in the oxidative stress response. *PLoS One* **6**, e27371.
66. Cino, E. A., Choy, W. Y. & Karttunen, M. (2012). Comparison of Secondary Structure Formation Using 10 Different Force Fields in Microsecond Molecular Dynamics Simulations. *J Chem Theory Comput* **8**, 2725-2740.
67. Fuxreiter, M. (2012). Fuzziness: linking regulation to protein dynamics. *Mol Biosyst* **8**, 168-177.
68. Fuxreiter, M. & Tompa, P. (2012). Fuzzy complexes: a more stochastic view of protein function. *Adv Exp Med Biol* **725**, 1-14.
69. Zhuo, Y., Ilangoan, U., Schirf, V., Demeler, B., Sousa, R., Hinck, A. P. & Lafer, E. M. (2010). Dynamic interactions between clathrin and locally structured elements in a disordered protein mediate clathrin lattice assembly. *J Mol Biol* **404**, 274-290.
70. Hammoudeh, D. I., Follis, A. V., Prochownik, E. V. & Metallo, S. J. (2009). Multiple independent binding sites for small-molecule inhibitors on the oncoprotein c-Myc. *J Am Chem Soc* **131**, 7390-7401.
71. Uversky, V. N. (2012). Intrinsically disordered proteins and novel strategies for drug discovery. *Expert Opin Drug Discov* **7**, 475-488.
72. Wang, J., Cao, Z., Zhao, L. & Li, S. (2011). Novel Strategies for Drug Discovery Based on Intrinsically Disordered Proteins (IDPs). *Int J Mol Sci* **12**, 3205-3219.

73. Cheng, Y., LeGall, T., Oldfield, C. J., Mueller, J. P., Van, Y. Y., Romero, P., Cortese, M. S., Uversky, V. N. & Dunker, A. K. (2006). Rational drug design via intrinsically disordered protein. *Trends Biotechnol* **24**, 435-442.
74. Cino, E., Fan, J., Yang, D. & Choy, W. Y. (2012). (1)H, (15)N and (13)C backbone resonance assignments of the Kelch domain of mouse Keap1. *Biomol NMR Assign.* In press.
75. Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K. & Uversky, V. N. (2010). PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim Biophys Acta* **1804**, 996-1010.
76. Ishida, T. & Kinoshita, K. (2008). Prediction of disordered regions in proteins based on the meta approach. *Bioinformatics* **24**, 1344-1348.
77. Yi, S., Brickenden, A. & Choy, W. Y. (2008). A new protocol for high-yield purification of recombinant human prothymosin alpha expressed in Escherichia coli for NMR studies. *Protein Expr Purif* **57**, 1-8.
78. Dosztanyi, Z., Meszaros, B. & Simon, I. (2009). ANCHOR: web server for predicting protein binding regions in disordered proteins. *Bioinformatics* **25**, 2745-2746.
79. Disfani, F. M., Hsu, W. L., Mizianty, M. J., Oldfield, C. J., Xue, B., Dunker, A. K., Uversky, V. N. & Kurgan, L. (2012). MoRFpred, a computational tool for sequence-based prediction and characterization of short disorder-to-order transitioning binding regions in proteins. *Bioinformatics* **28**, i75-83.
80. Marsh, J. A., Singh, V. K., Jia, Z. & Forman-Kay, J. D. (2006). Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. *Protein Sci* **15**, 2795-2804.
81. Kay, L. E. (1998). Protein dynamics from NMR. *Nat Struct Biol* **5 Suppl**, 513-517.
82. Walma, T., Aelen, J., Nabuurs, S. B., Oostendorp, M., van den Berk, L., Hendriks, W. & Vuister, G. W. (2004). A closed binding pocket and global destabilization modify the binding properties of an alternatively spliced form of the second PDZ domain of PTP-BL. *Structure* **12**, 11-20.
83. Peng, J. W. & Wagner, G. (1992). Mapping of the spectral densities of N-H bond motions in eglin c using heteronuclear relaxation experiments. *Biochemistry* **31**, 8571-8586.
84. Fan, J. S., Lim, J., Yu, B. & Yang, D. (2011). Measurement of amide hydrogen exchange rates with the use of radiation damping. *J Biomol NMR* **51**, 151-162.
85. Radford, S. E., Buck, M., Topping, K. D., Dobson, C. M. & Evans, P. A. (1992). Hydrogen exchange in native and denatured states of hen egg-white lysozyme. *Proteins* **14**, 237-248.
86. Kim, J. H., Fuzery, A. K., Tonelli, M., Ta, D. T., Westler, W. M., Vickery, L. E. & Markley, J. L. (2009). Structure and dynamics of the iron-sulfur cluster assembly scaffold protein IscU and its interaction with the cochaperone HscB. *Biochemistry* **48**, 6062-6071.
87. van Ingen, H., Baltussen, M. A., Aelen, J. & Vuister, G. W. (2006). Role of structural and dynamical plasticity in Sin3: the free PAH2 domain is a folded module in mSin3B. *J Mol Biol* **358**, 485-497.

88. Komatsu, M., Kurokawa, H., Waguri, S., Taguchi, K., Kobayashi, A., Ichimura, Y., Sou, Y. S., Ueno, I., Sakamoto, A., Tong, K. I., Kim, M., Nishito, Y., Iemura, S., Natsume, T., Ueno, T., Kominami, E., Motohashi, H., Tanaka, K. & Yamamoto, M. (2010). The selective autophagy substrate p62 activates the stress responsive transcription factor Nrf2 through inactivation of Keap1. *Nat Cell Biol* **12**, 213-223.
89. Hutchinson, E. G. & Thornton, J. M. (1994). A revised set of potentials for beta-turn formation in proteins. *Protein Sci* **3**, 2207-2216.
90. Lau, A., Wang, X. J., Zhao, F., Villeneuve, N. F., Wu, T., Jiang, T., Sun, Z., White, E. & Zhang, D. D. (2010). A noncanonical mechanism of Nrf2 activation by autophagy deficiency: direct interaction between Keap1 and p62. *Mol Cell Biol* **30**, 3275-3285.
91. Mittag, T., Orlicky, S., Choy, W. Y., Tang, X., Lin, H., Sicheri, F., Kay, L. E., Tyers, M. & Forman-Kay, J. D. (2008). Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc Natl Acad Sci US A* **105**, 17772-17777.
92. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L. & Sykes, B. D. (1995). ¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR. *J Biomol NMR* **6**, 135-140.
93. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293.
94. Johnson, B. A. (2004). Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* **278**, 313-352.
95. Connelly, G. P., Bai, Y., Jeng, M. F. & Englander, S. W. (1993). Isotope effects in peptide group hydrogen exchange. *Proteins* **17**, 87-92.
96. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins* **17**, 75-86.
97. Hess, B., Kutzner, C., Spoel, D. (2008). Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* **4**, 435-447.
98. Oostenbrink, C., Soares, T. A., van der Vegt, N. F. & van Gunsteren, W. F. (2005). Validation of the 53A6 GROMOS force field. *Eur Biophys J* **34**, 273-284.
99. Oostenbrink, C., Villa, A., Mark, A. E. & van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* **25**, 1656-1676.
100. Brunger, A. T. (2007). Version 1.2 of the Crystallography and NMR system. *Nat Protoc* **2**, 2728-2733.
101. MacArthur, M. W. & Thornton, J. M. (1993). Conformational analysis of protein structures derived from NMR data. *Proteins* **17**, 232-251.
102. Bussi, G., Donadio, D. & Parrinello, M. (2007). Canonical sampling through velocity rescaling. *J Chem Phys* **126**, 014101.
103. Parrinello, M. & Rahman, A. (1981). Polymorphic transitions in single crystals: A new molecular dynamics method *J Appl Phys* **52**, 7182-7190.

104. Hess B, B. H., Berendsen HJC, Johannes JGEM. (1997). A linear constraint solver for molecular simulations. *J Comput Chem* **18**, 1463-1472.
105. Darden T, Y. D., Pedersen L. (1993). Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems *J Chem Phys* **98**, 10089-10093.

Figure Legends

Figure 1. NMR titration experiments of ProTα with Kelch. (a) ^1H - ^{15}N HSQC spectra of ProTα alone (blue) and with 2 equivalents of Kelch domain (red). Close-up views of the ^1H - ^{15}N HSQC spectra, showing selected chemical shift changes during the titration, are shown. (b) SSP scores for ProTα in the free (blue) and bound state with Kelch (red), calculated using $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ chemical shifts. Residues for which SSP scores were not calculated are represented by a tick mark. The regions showing significant changes in SSP correspond to peaks on the HSQC spectrum that undergo large chemical shift changes.

Figure 2. Backbone ^{15}N relaxation experiments of ProTα in the free (blue) and bound state with 2 equivalents of Kelch (red). (a) R_1 (longitudinal) relaxation rate. (b) R_2 (transverse) relaxation rate. (c) steady-state ^1H - ^{15}N NOE.

Figure 3. Protection factor (P) of backbone amides in ProTα in the absence (blue bars) and presence of Kelch (red *). The P value of each amino acid was calculated as the ratio of the intrinsic HX rate (k_{int}) of random coil to the experimentally determined HX rate. The k_{int} values were estimated using the SPHERE web server (<http://www.fccc.edu/research/labs/roder/sphere/sphere.html>)^{95,96}

Figure 4. Permutation peptide array experiments. 20-mer peptides of ProTα (residues 38-57) were immobilized on a cellulose membrane and then probed with the Kelch domain of Keap1. The left column shows the wild-type sequence and positions of the amino acids that were permuted (mutations specified at the top of second column). Cys was excluded in the permutation to avoid false positives. The '+' in the second column indicates wt 20-mer sequence and is used as a positive control. Array was repeated twice. Intensity of each permutation was qualitatively measured and compared to wild-type, and results reported are based on what was observed consistently in the two arrays. Columns, on the right side indicate amino acid residues that weaken, strengthen, or have no effect on the binding affinity between the 20-mers and Kelch. As shown in the array, mutations

in the EENGINE motif are poorly tolerated. AA, amino acid; X, every amino acid except Cys and those mentioned as weaker or stronger; wt, wild-type.

Figure 5. ITC profiles of Keap1 with full length ProTα and its mutant variants. (a) Full length wild-type ProTα (b) EENGINE ProTα 11-mer peptide (c) N36A (d) E41P (e) G44A (f) E45G (g) E47A (h) E47R (i) E51G

Figure 6. Average circular variances from the Kelch-bound state of ProTα and of Nrf2 MD simulations. The values were calculated as described by¹⁰¹ and averaged over the last 200 ns (100-300 ns) of the trajectories. Values for the Kelch domain binding domain and surrounding residues are shown.

Figure 7. Backbone rmsd's between the bound state and MD structures. The rmsd values were calculated by least-squares fitting the backbone atoms (N, Cα, and C) of the NEEN or NPEN sequences from each frame to the corresponding atoms of bound state reference structure (PDB id: [2Z32](#))⁵².

Figure 1

Figure 2

Figure 4

AA Position (wt)	Amino acid substituted in ProTα and probed with Kelch	Effect of substitution on Kelch binding			Essential for binding
		Weaker	Equal	Stronger	
N38		X	KRFYW		N
A39		HGPA	ILVMFYSW	KRDEQT	A
N40		X	HN		N
E41		NSTGVMYW	X	P	E
E42		X	E		E
N43		X	NH	ST	N
G44		X	G		G
E45		X	E		E
Q46		Y	X		-
E47		X	RDEW		-
A48		Y	X		-
D49		Y	X		-
N50		Y	X		-
E51		Y	X		-
V52		Y	X		-
D53		Y	X		-
E54		Y	X		-
E55		Y	X		-
E56		X	X		-
E57		X	X		-

Figure 5

Figure 6

Figure 7

Table 1. Thermodynamic parameters for binding of the full-length ProTα, ProTα peptide, and ProTα mutants to Kelch domain of Keap1 at 25 °C*

WT Kelch	n	K_a (10^5 M^{-1})	K_d (μM)	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)	ΔG (kcal/mol)
Full length ProTα	1.06 ± 0.02	3.8 ± 0.6	2.6 ± 0.4	-14.8 ± 0.4	-7.19	-7.61 ± 0.09
ProTα Peptide EENGE	$0.93 \pm 0.00(4)$	5.6 ± 0.3	1.79 ± 0.09	-22.6 ± 0.1	-14.76	-7.84 ± 0.03
N36A	$0.95 \pm 0.00(5)$	1.53 ± 0.05	6.5 ± 0.2	-10.1 ± 0.6	-3.03	-7.07 ± 0.02
E41P	$1.10 \pm 0.00(2)$	19.2 ± 0.8	0.52 ± 0.02	-14.17 ± 0.04	-5.60	-8.57 ± 0.02
E47A	1.08 ± 0.16	0.069 ± 0.006	145 ± 13	-15.2 ± 0.3	-9.96	-5.24 ± 0.05
E47R	0.93 ± 0.09	0.33 ± 0.03	30 ± 3	-13.1 ± 0.2	-6.94	-6.16 ± 0.05
D49A	1.04 ± 0.02	1.5 ± 0.2	6.7 ± 0.9	-12.6 ± 0.3	-5.53	-7.07 ± 0.08
E51A	1.06 ± 0.01	1.14 ± 0.05	8.8 ± 0.4	-12.7 ± 0.2	-5.80	-6.90 ± 0.03
E51G	1.04 ± 0.01	1.40 ± 0.06	7.1 ± 0.3	-15.23 ± 0.02	-8.21	-7.02 ± 0.02

* n is the stoichiometry, K_a , and K_d are the association and dissociation constants, respectively. $\Delta G = -RT \ln K_a$, where T is the temperature in Kelvin, R is the gas constant. ΔH and ΔS are the enthalpy and entropy changes, respectively. Each experiment was repeated twice; the values obtained between the two runs were similar. Values listed in the tables are based on one run, and the data for the duplicate run is provided in the supplementary data (Table S2).

Supplemental Data

Figure S1. Predictions of disorder propensity and potential binding sites for ProT α . POND-R-FIT⁷⁵ (red) and metaPrDOS⁷⁶ (green) plots analyzing disorder nature of ProT α . ANCHOR⁷⁸ and MoRFpred⁷⁹ analysis of potential binding sites in ProT α . The blue line represents the predicted binding site by ANCHOR, and the pink line represents the predicted binding interface by MoRFpred.

Figure S2. ^1H – ^{15}N TROSY-HSQC spectra of the human Kelch domain of Keap1 (black). Overlay of the Kelch spectra in the presence of the full-length human ProTa (red) and ProTa peptide (cyan), $^{33}\text{APANGNANEENGQEADNEV}^{52}$.

Figure S3. Sedimentation equilibrium analysis of Kelch and ProT α binding. Pure Kelch protein at a concentration of 10 μ M was centrifuged with various concentrations of ProT α (to achieve a ratio of 1:0.5, 1:1, and 1:2 with the Kelch domain) at 4°C at 22 000 rpm. The upper panel represents the residual values for the fits.

Figure S4. Snapshots from the A) Kelch-ProT α and B) Kelch-Neh2 domain MD simulations. Conformations captured at $t=50$, 100, 150, 200, 250 and 300 ns are shown. The regions correspond to the bound-state β -turn structure are circled. Cluster center structures illustrate molecular detail of the binding interfaces.

Figure S5. Snapshots from the free state MD simulations of the A) wild type ProT α peptide and B) E41P ProT α peptide.

Table S1. Molecular weight values of Kelch calculated from the sedimentation equilibrium experiment.

Protein	Concentration (μM)	Calculated MW (Da)	Observed MW (Da)
Kelch	10	33915.85	34565 $\pm$ 338
Kelch	10	39952.76	41551 $\pm$ 331
ProT α	5		
Kelch	10	45989.67	44978 $\pm$ 396
ProT α	10		
Kelch	10	45989.67	46793 $\pm$ 379
ProT α	20		

Table S2. Thermodynamic parameters for binding of the ProT α peptide, full-length ProT α and its mutants to Kelch domain of Keap1 at 25 °C (second set)

WT Kelch	n	K_a (10^5 M^{-1})	K_d (μM)	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)	ΔG (kcal/mol)
Full length ProT α	1.04 ± 0.02	3.3 ± 0.3	3.04 ± 0.2	-15.5 ± 0.4	-7.97	-7.53 ± 0.05
ProT α Peptide EENGINE	1.13 ± 0.004	5.7 ± 0.3	1.74 ± 0.08	-23.1 ± 0.1	-15.24	-7.86 ± 0.03
N36A	0.95 ± 0.04	1.6 ± 0.5	6.1 ± 1.8	-9.9 ± -0.5	-2.8	7.1 ± 0.2
E41P	1.16 ± 0.002	20.3 ± 0.8	0.49 ± 0.02	-15.05 ± 0.05	-6.45	-8.60 ± 0.02
E47A	1.10 ± 0.12	0.061 ± 0.002	164 ± 5	-15.9 ± 0.2	-10.74	-5.16 ± 0.02
E47R	0.91 ± 0.04	0.303 ± 0.02	33 ± 2	-13.20 ± 0.08	-7.09	-6.11 ± 0.03
D49A	1.06 ± 0.05	2.2 ± 0.8	4.5 ± 1.6	-12.0 ± 0.7	-4.7	-7.3 ± 0.2
E51A	0.95 ± 0.01	1.18 ± 0.06	8.5 ± 0.4	-13.2 ± 0.2	-6.28	-6.92 ± 0.03
E51G	1.05 ± 0.01	1.49 ± 0.05	6.7 ± 0.2	-15.63 ± 0.02	-8.57	-7.06 ± 0.02

* n is the stoichiometry, K_a , and K_d are the association and dissociation constants, respectively.

$\Delta G = -RT \ln K_a = \Delta H - T\Delta S$, where T is the temperature and R is the gas constant. ΔH and ΔS are the enthalpy and entropy changes, respectively.

Materials and Methods

Analytical ultracentrifugation

Analytical ultracentrifugation technique was employed to determine the binding stoichiometry of ProTα to the Kelch domain of Keap1. Sedimentation equilibrium experiments were performed in a Beckman-Coulter Optima XL-A analytical ultracentrifuge (Beckman Coulter, Palo Alto, CA) with an AN60Ti rotor. Six-sector Epon-charcoal centerpieces with quartz windows were used. Protein samples were dialyzed in 50 mM phosphate, 100 mM NaCl, 1 mM DTT, at pH 7.0. Data were collected at 22,000 rpm following 24 h equilibration period at 4°C. Kelch absorbance at 280 nm was measured with 10 replicates with a step size of 0.002 cm. Data were fitted to a single species model using Microsoft Origin 7.0 software. Partial specific volume for Kelch was calculated to be 0.7219 mL/g based on the amino acid composition. The solvent density was measured to be 1.008 g/mL. Protein concentrations were calculated by protein quantification methods: Lowry for ProTα and Bradford assay for Kelch, as well as by amino acid analysis (Advanced Protein Technology Centre, The Hospital for Sick Children, Toronto, ON).

Primers used for point mutations.

E41P: 5' CCCTGCTAACGGGAATGCTAATCCGGAAAATGGGGAGC 3'

E45G: 5' TAATGAGGAAAATGGGGGCGCAGGAGGCTGACAATG 3'

E45A: 5' TAATGAGGAAAATGGGGGCGCAGGAGGCTGACAATG 3'

E51A: 5' CAGGAGGCTGACAATGCGGTAGACGAAGAAGAG 3'

E51G: 5' CAGGAGGCTGACAATGGGGTAGACGAAGAAGAG 3'

E47A: 5' GAAAATGGGGAGCAGGCGGCTGACAATGAGGTA 3'

E47R: 5' GAGGAAAATGGGGAGCAGCGCGCTGACAATGAGGTAGAC 3'

D49A: 5' GGAGCAGGAGGCTGCCAATGAGGTAGACG 3'

N36A: 5' ATGGAAGAGACGCCCCTGCTGCCGGAATGCTAATG 3'

G44A: 5' GGAATGCTAATGAGGAAAATGCGGAGCAGGAGGCTG 3'

MD Simulations

Generation of bound-state complexes

The ProTα coordinates were obtained from a previously reported free state simulation in which bound-state-like hairpin structures were observed⁵². This step allowed for straightforward modeling of the full-length protein onto the existing crystal structure coordinates⁵². For the Neh2 domain, we used the Crystallography & NMR System (CNS)¹⁰⁰ to generate an extended structure based solely upon amino acid sequence and performed simulated annealing simulations in which dihedral angle restraints were applied to force residues 39-71 into a helical conformation⁴⁷. A structure from the annealing simulations was pre-equilibrated using unrestrained MD in explicit solvent for several hundred ns. During this simulation, hairpin formation in the Kelch binding region occurred, and the central helix remained largely intact. We proceeded to model the Neh2 domain coordinates from this system onto the existing crystal structure coordinates⁵⁴. When modeling the complexes, the protein coordinates from the crystal structures^{52,54} were used in their complete and unmodified forms, with the exception of the addition of N- and C-terminal acetyl (ACE) and NH₂ capping groups, respectively, to

the Kelch and Neh2 domains. The bound-state complexes were pre-equilibrated for 130 ns each in explicit solvent before production runs. Both the ProTα-Kelch domain and Nrf2-Kelch domain complexes remained intact (did not dissociate) throughout equilibration.

Parameters for MD simulations

After a 130 ns equilibration of each system in explicit solvent, a production simulation was performed. The protonation states of ionizable amino acids were chosen based upon their most probable charge at pH 7. Cys side-chains were always protonated, and His side-chain protonation states were determined automatically by GROMACS⁹⁷, based upon their local hydrogen bonding geometry. Cubic boxes with side lengths of 11 and 10 nm were used for the ProTα and Neh2 systems, respectively. Water molecules were replaced with sodium (Na⁺) and chloride (Cl⁻) ions as necessary to neutralize the system charges and reach a physiological salt concentration to 0.1 M. The ProTα-Kelch domain system contained 41812 water molecules, 130 Na⁺ and 80 Cl⁻. The Neh2 domain-Kelch domain system contained 31022 water molecules, 78 Na⁺ and 60 Cl⁻. The protein and non-protein atoms were coupled to separate temperature baths, maintained at 310 K using the Parrinello-Bussi algorithm¹⁰². The pressure was maintained isotropically at 1 bar using the Parrinello-Rahman barostat¹⁰³. A 2-fs timestep was employed. Periodic boundary conditions were applied in all directions. Before production runs, the energy of each system was minimized using the steepest descents algorithm, followed by 2 ps of position-restrained dynamics with all non-hydrogen atoms restrained with a 1000 kJ mol⁻¹ force constant. The atom velocities were assigned from a Maxwellian distribution at 310

K. Bond lengths were constrained using the LINCS algorithm¹⁰⁴. The cut-off for Lennard-Jones interactions was 1.0 nm. Dispersion corrections for energy and pressure were applied. The Particle-Mesh Ewald (PME) method¹⁰⁵, with 0.12 nm grid-spacing and a 1.0 nm real-space cut-off, was used for electrostatics. The simulations were run for 300 ns each and the last 200 ns were used for circular variance analysis.