**“Physical basis for modulation of folding state in a disordered protein”**

**Abstract**

The human von Hippel-Lindau (VHL) tumor suppressor is protein with low thermodynamic stability and a large disordered region that has been studied successfully as a model substrate of refolding and degradation pathways in the cell. It is thought to interact with different chaperones and requires binding to cofactors like elongin C to become stable in the crowded intracellular environment, such that cell stress is likely to lead to a misfolded version of the protein. When this gene is transfected into yeast, the resulting protein is quickly degraded by the cell’s quality control machinery. Here, we combine computational methods with *in vivo* experiment to study the basis in amino acid sequence for the misfolding propensity of VHL. Using a previously developed “burial mode” model of the trade-off between hydrophobic and steric effects in the protein globule, we generated a set of 20 mutant versions of the human VHL protein created by swapping two residues (‘pair-swap mutations’) that were transfected into *S. cerevisiae* cells and tagged fluorescently. Of this group, two mutations produced proteins that were stable enough to be detected in large quantities by FACS analysis. One mutation directly affected a known chaperonin binding site (V155<->G19), while the other mutation (L201<->E173) is not in a known interaction region. Subsequent modeling revealed that the L201E173 mutant was predicted to have conformational changes that affect co-factor and chaperone interaction. These findings suggest that the burial mode model can help elucidate mechanisms of protein misfolding and cellular persistence through its predictions of conformational changes.

**Introduction**

To function properly in the cell, a globular protein chain typically must remain folded into a specific conformation or set of conformations known as its native state. A primary determinant of how a globular protein folds is its amino acid sequence, which fixes the pattern of internal and external forces that act on the polypeptide chain in the aqueous environment. When a protein either cannot reach or cannot maintain its native conformation in the cell it is considered to be in a misfolded state, which is often accompanied by a loss of cellular function. Misfolded proteins also exhibit a marked tendency to associate non-specifically, and sometimes form potentially cytotoxic aggregates if left in the crowded intracellular environment (Chiti and Dobson, 2006, 2009; Luby-Phelps, 2000), and a large range of globular proteins have been shown to be capable of forming amyloid fibrils under partially-denaturing conditions (Chiti and Dobson, 2009). Mutations that lead to protein misfolding and/or aggregation have been implicated in proteopathies such as Huntington’s and the prion-based Creutzfeldt-Jakob disease, emphasizing a need to better understand the cellular response to protein misfolding in the context of the physical driving forces that govern how an amino acid sequence can reach its native structure.

The protein quality control (PQC) machinery, found in different forms across all kingdoms of life, consists of the cellular pathways linked to protein folding and misfolding. Chaperone proteins, a key component of PQC systems, can either assist a protein in folding or can target incorrect conformations for either refolding or destruction through the ubiquitin proteasome pathway or autophagy (Kim et al., 2013). Despite the prevalence of PQC machinery across organisms, many aspects of the system are not well understood. Misfolded proteins are able to reach a wide variety of different non-native states, and the PQC must be able to target this diverse group of conformations either to assist refolding or to target them for destruction before the misfolded conformation can cause detrimental effects to the cell. Understanding how PQC systems recognize these misfolded substrates is an ongoing area of investigation. Hsp70, a heat-shock induced chaperone found in both bacteria and eukaryotes, has been shown to recognize exposed hydrophobic sites, particularly short (~5-7 residue) hydrophobic sequences flanked by positively-charged amino acids (Marcinowski et al., 2013; Rüdiger et al., 1997). The eukaryotic chaperonin TRiC, on the other hand, has eight distinct subunits that are each capable of recognizing distinct motifs in a variety of substrates, with mutations in different subunits leading to different cellular phenotypes (Amit et al., 2010; Spiess et al., 2006). The need to control protein conformations is also present in the nucleus, with a recent discovery suggesting that the ubiquitin ligase nuclear protein San1 can recognize exposed hydrophobicity and substrate protein insolubility and target these aberrant proteins for destruction in the nucleus of the cell through the use of multiple points of contact with San1 binding sites interspersed in a highly flexible domain (Rosenbaum et al., 2011). How the PQC recognizes the full range of malformed proteins, however, remains an ongoing question.

Marginally stable proteins, which can be induced to misfold easily and exist at a ‘tipping point’ between stable well-folded conformations and PQC-targeted misfolded variants, have been used as an experimental mechanism to explore PQC substrate recognition and subsequent refolding and degradation pathways. The human von Hippel-Lindau protein (pVHL) is one such example that is particularly susceptible to incorrect folding. This model misfolding protein forms part of an E3 ubiquitin ligase complex that targets molecules like HIF-1α for degradation, and has been catalogued in depth because hundreds of different mutations have been linked to cancer pathways in humans (Nordstrom-O’Brien et al., 2010). The first ~60 residues of the 213-residue protein remain disordered in the protein’s native state, and it must undergo a distinct folding pathway *in vivo* including interacting with chaperones such as Hsp70 and TRiC and binding with its cofactors elongin B and elongin C to achieve a state resistant to cellular degradation (McClellan et al., 2005; Schoenfeld et al., 2000). For TRiC, two short motifs – Box 1 and Box 2 – have been shown to be necessary and sufficient for TRiC binding to pVHL in yeast (Feldman et al., 2003). When folded correctly in complex with its cofactors, the non-disordered region adopts a well-defined tertiary structure; however, the protein adopts a molten globule state without its binding partners *in vitro* that consists of a partially collapsed state with some secondary structure but no tertiary structure (Sutovsky, 2004). This molten globule state indicates that pVHL has difficulty achieving its native state without interactions with other proteins. Perturbations to the system, including mutations to pVHL, are likely to lead to a misfolded or otherwise nonfunctional version of the protein *in vivo* (Hansen et al., 2002; Knauth et al., 2006). When VHL is introduced into non-native systems like *S. cerevisiae* or *E. coli*, where it does not exist naturally, the protein cannot achieve a biologically stable state and in yeast is quickly degraded by the cell (Melville et al., 2003; Sutovsky, 2004). In the context of understanding PQC in yeast, pVHL is a protein whose typical state is poised between adequate folding and being targeted for destruction through a misfolding recognition pathway.

One of the persistent difficulties in understanding the physical mechanisms of protein misfolding and bases for recognition of substrates by the PQC is that almost by definition, misfolded proteins are not amenable to conventional methods of structural characterization. Protein chains that adopt main different conformations cannot easily be crystalized, and aggregation-protein proteins are difficult to solubilize for *in vitro* characterization. Thus, in examining the effects of different mutations on a marginally stable protein like pVHL, a computational model that could give insight into the resulting structural changes could offer a new and much needed perspective on the connection between sequence, structure, and recognition by PQC machinery for a large number of sequences. Recently, a phenomenological model developed by J.L. England to predict tertiary structural information from sequence alone in globular proteins, has shown promise as a method of computationally exploring the allowed conformational space of fluctuating protein folds (England, 2011). The burial trace is computed by minimizing an energy function consisting of the hydropathies of each residue and the stretching between neighbor amino acids, subject to steric constraints. The calculation generally takes less than a second to run for short sequences, and adding noise to the parameters of the system can generate an ensemble of amino acid burial patterns for a given protein sequence that can be used to investigate the variability in structures that a protein can adopt. The rapidity of this model in determining structural information makes it an excellent candidate for probing large numbers of potential mutations of marginally stable proteins to understand PQC substrate recognition *in silico* and to guide *in vivo* experiments.

To investigate the link between the underlying biophysics of protein folding and the role of the PQC machinery in recognizing and clearing misfolded proteins, the burial mode model was used to investigate the folding characteristics of the human von Hippel-Lindau tumor suppressor protein. Burial traces were calculated to predict exposed residues for the lowest energy conformations of different mutations of pVHL, 20 of which were generated experimentally and tested for the degradation properties *in vivo*. Our findings support the use of burial mode analysis as a powerful new tool for elucidating the structural basis of misfolding and QC recognition in eukaryotes.

**Results**

Our initial goal was to see if we could use computational modeling of the VHL protein to design mutations that would alter its misfolding and degradation. Since VHL is a human protein, it lacks its usual elongin co-factors when expressed in yeast, and thus is quickly degraded. We began by using burial mode analysis to design mutations of VHL that would be expected to reduce the protein’s misfolding propensity, and thus slow the degradation of the protein in the yeast cytosol..

*Description of Biophysical Model*

A previous work by one of the co-authors presented a phenomenological framework to estimate the “burial trace,” or the distance of each amino acid from the center of mass in the lowest-energy fold of a globular polypeptide chain.  This burial trace predictor assumes that each amino acid behaves like it is connected by a spring to its neighbor residues, simulating peptide bonds (Fig. 1).  These hypothetical springs contribute to the overall energy for the system, along with the energetically unfavorable terms of putting hydrophilic residues far from the center of the protein and burying hydrophobic ones.  This energy function can then be minimized under spatial constraints, specifically to ensure that the individual peptides are not clustered into a small volume that would lead to prohibitive steric clashing in a physical scenario. Therefore, the simulated fold of the peptide chain is constrained to have each residue’s average squared distance from the center of mass of the protein be 3R2/5, derived from assuming the protein adopts a roughly spherical form. A linear solver can then be used to solve for a minimal folding energy along with the burial trace. The calculation generally takes less than a second to run for short sequences, and is especially appropriate for small alpha-helical globular proteins. The resulting calculation gives a prediction of the distance of each residue from the center of the protein in its lowest energy state – the burial trace.

VHL’s relatively small size (213 residues) and predominance of alpha helical structure in its non-disordered region make it a good candidate for use of the burial mode model. As a first attempt at modeling VHL stability, all possible VHL pair-swap mutations (where amino acids at two different locations in the sequence are switched, which preserves overall residue frequencies for the sequence) were ranked according to a parameter called structural variability. This parameter simulates generating an ensemble of possible conformers at an energy level slightly above the ground state for each mutated sequence using the burial trace model, with the underlying hypothesis that sequences that have very large differences in allowable burial traces might not have a clear energy minimum that facilitates stable folding. The number of available conformations that are close to the ground state, as modeled by structural variability in this model, can then be related to thermodynamic stability in terms of how likely the protein is to be found in a burial trace close to its native state at low energies. All possible pair-swap mutations were ranked by low structural variability scores, and ten low-scoring mutants were chosen to be created experimentally along with a control group of ten randomly-chosen mutants that did not demonstrate distinctive structural variability, with the mutations listed in Table 1.

|  |  |
| --- | --- |
| Low Structural Variability Mutations | Random Mutations |
| 1. K171 P61 2. R176 G14 3. V20 R3 4. P99 V74 5. V66 P40 6. K171 P5 7. V20 R4 8. R167 A11 9. C162 N7 10. Q203 N109 | 1. E173 K196 2. E70 G19 3. V155 G19 4. T124 K159 5. T157 V194 6. D28 A56 7. G93 V181 8. R69 S168 9. L201 E173 10. S183 C162 |

**Table 1***. List of mutants that were created experimentally, half of which were chosen based on a model parameter and half of which were chosen randomly.*

All 20 mutations were then created in an *S. cerevisiae* constitutive plasmid with an attached fluorescent Dendra tag and transfected into budding yeast cells, which do not contain native VHL protein, nor the human elongin co-factors. If the VHL protein can exhibit cellular persistence in these cells, then the attached fluorescent tag would also be preserved and observable. Fluorescence was detected by using Fluorescence-Activated Cell Sorting (FACS analysis), with results shown in Fig. 2 for most mutants. Interestingly, the first ten non-randomly chosen mutations exhibited lower overall fluorescence (and therefore less cellular persistence) than the ten randomly-chosen control mutations. A maximum-likelihood estimate between the random and non-random sets bounded the probability of observing no highly fluorescent hits in the non-random set while obtaining two or more hits in the random control set at a maximum value slightly below 0.1. It is possible that the structural variability calculations actually predicted the opposite effect than originally thought – namely, lower cellular persistence rather than higher thermodynamic stability – perhaps due to kinetic trapping in conformations that are more subject to the cell’s PQC degradation pathways.

Furthermore, two mutations in the random control set achieved greater than 60% of the fluorescence level of the individual Dendra tag control. The V155-G19 pair swap mutation, which demonstrated ~64% of the fluorescence of the Dendra tag alone, changed a hydrophobic residue in the Box 2 region (spanning residues 148-155) which mediates the chaperonin TRiC’s interaction with VHL and replaced it with the smaller and more flexible glycine, which is normally located in the disordered N-terminal region. This mutation was distinct in our set because it was the only one that directly affected a known chaperonin binding site.

This observation suggested that a possible explanation for the stabilizing mutant could be related to how different chaperone binding sites are affected. The L201-E173 mutation with the highest level of fluorescence exhibited ~75% of the baseline Dendra level, which corresponds to a five-fold increase in the protein’s ability to escape PQC degradation. The biophysical characteristics of the residues swapped were dissimilar, with the hydrophobic leucine at residue 201 changing into a negatively charged glutamic acid (normally at residue 173) and vice-versa. Although both mutations occurred in the region that adopts a well-defined conformation during correct folding, neither was in a known binding site for cofactors or chaperones, raising the possibility that an allosteric response could explain the structural change necessary to evade PQC degradation systems.

Past work has established several interaction sites in the well-folded region of VHL, including two short motifs called Box 1 (residues 114-118) and Box 2 (residues 148-155) that are known to bind to chaperonins (Feldman et al., 2003) and a short binding motif for pVHL’s cofactors elongin B and elongin C. A BLAST search for pVHL cofactors elongin B and elongin C also revealed that although elongin B does not have a counterpart in yeast, a homolog of elongin C is present. Furthermore, this homolog has been shown to bind to a pVHL fragment containing residues from 157-171, which contains the known elongin B/C binding site, *in vitro* (Botuyan et al., 2001). Since elongin B and C stabilize pVHL in human cells, one testable hypothesis is that burying the chaperonin interaction sites (and therefore possibly protecting pVHL from chaperone-assisted degradation) and exposing the elongin interaction site could help explain why certain mutations caused the protein to persist longer in the cell.

To explore how mutations could affect conformations, particularly in understanding exposure patterns of different interaction sites, the burial trace of each experimentally-created mutant was computed using the England model and compared to the actual burial trace calculated from the crystal structure of the well-folded region. Since the basic burial trace model works best on alpha-helical structures, each input sequence was truncated to residues 63-204 corresponding to the region that can be crystallized (PDB 1VCB) in an attempt to improve accuracy within the region that can adopt a well-defined folding state (Stebbins et al., 1999). Fig. 3 shows the predicted burial traces for the wildtype and L201-E173 sequences, along with maximal and minimal predicted burial values at each residue for all 20 experimentally-created mutations. The experimental burial trace obtained directly from the crystal structure of wildtype VHL (PDB 1VCB chain C) is also shown. The crystal structure burial trace shows a distinctive burial of the Box 1 and Box 2 regions, suggesting that TRiC helps to bury these motifs in the native conformation. The elongin binding site is also moderately exposed compared to Box 1 and Box 2.

Out of all the experimentally-created mutants, the well-stabilized L201-E173 mutation (black line) is predicted to have the most buried Box 1 and Box 2 regions, similar to how TRiC is able to bury these motifs in achieving its native conformation in human cells. Furthermore, the elongin interaction region of the L201-E173 mutation was distinctive in that the model predicted that it was one of the most exposed regions in the entire sequence. This property of burying known chaperone interaction sites and exposing the cofactor binding site is unique compared to the predicted results of the wildtype protein (orange line) as well as the other experimentally-created mutants (grey envelope), and is more similar to the known native conformation of the protein derived from its crystal structure. To study the significance of this finding, we also calculated each possible pair swap mutation in the region of VHL that can be crystallized (residues 63-204) and compared their burial traces by Pearson correlation to the crystal structure’s burial trace. Only 6.18% of possible mutations in this region had a higher correlation with the crystal structure than the L201-E173 mutant (Fig 4a).

The two stabilizing mutants were unique in that they affected chaperone binding sites, either through direct mutation as in V155-G19 or through causing the burial of these sites to be more energetically favorable. The other distinctive component of the most stable mutant’s burial trace, however, was its extreme exposure of its elongin cofactor binding site. Although the crystal structure moderately exposes this region, the relatively small distance from the center of the protein in this region could skew our understanding of how rare this burial pattern might be. Therefore, all possible pair swap mutations were also scored by ranking directly how well the Box 1 and Box 2 chaperonin interaction regions were buried and the elongin interaction site was exposed based on the burial model’s predictions. This score can be obtained from a mutated sequence by calculating the sum of the predicted distances of the most-buried residues in the Box 1 and 2 regions from the center of mass of the protein, and adding this subscore to the predicted distance of the maximally-exposed residue in the elongin interaction region from the most exposed residue of the entire protein as illustrated in Fig 4. The smallest score would correspond to Box 1 and Box 2 being located at the center of the protein and the elongin interaction site being on the most external part of the protein's surface, corresponding to an optimization of theoretically making the protein unrecognizable to PQC systems.

Each possible pair swap mutation, with the input truncated to the region [63,204], was scored according to this metric. Intriguingly, L201-E173 has the lowest (best) score out of ~10,000 possible mutations (Fig. 4b).  The location of the other experimentally-determined mutants is also shown, indicating a large gap in scores between mutant L201-E173 and the rest of the experimental sequences. The scores of the rest of the sequences, however, did not correlate with their fluorescence levels, indicating that this may be a binary effect.

Elongin C is not an essential protein in *S. cerevisiae*, so to test our hypothesis that the L201-E173 mutation helped stabilize the protein by decreasing chaperone interactions and/or interacting with yeast elongin C, the 20 mutated pVHL sequences were transfected into an elongin C knockout of yeast. The presence of fluorescence using quantitative cell sorting was again used to measure cellular persistence of the pVHL mutants, with results shown in Fig. 5. The stable expression of L201-E173 was decreased to that of the wildtype, indicating that elongin C could be helping to stabilize this mutation and lending credence to the findings of the burial model.

INSERT INFO ABOUT ANY NEW DATA

**Discussion**

The ability of a cell to recognize aberrant proteins and target them for either refolding or destruction is crucial to maintaining protein homeostasis and preventing pathological aggregation. To perform this function, proteins involved in PQC processes must be able to recognize a wide variety of possible misfolded conformations yet still be able to distinguish them from proteins that may have disordered regions that are part of their native conformations. Despite the importance of these systems in protecting the cell from misfolded protein accumulation, many questions remain about how PQC-related proteins can identify their targets.

A persistent challenge in studying how the PQC recognizes substrates is that substrates by nature tend to be structurally disordered and are therefore not easily amenable to traditional methods of *in vitro* structural characterization.  As a result, there is an opportunity for computational methods to play a central role in providing crucial structural information that can be used in combination with experimental results to better understand the workings of the PQC machinery. This study presents a novel method of integrating computational biophysics techniques to make predictions for how the PQC can recognize its substrates, and these predictions are then used to guide experimental results. This approach of using *in silico* results to guide *in vivo* experiments, and vice-versa, demonstrates that both approaches can be used in tandem to yield better explanations for physical phenomena.

In particular, pVHL has been used as a model substrate for chaperone proteins involved in assisted folding pathways as well as for studying PQC systems due to its marginal stability. Hsp90, Hsp70, and TRiC/CCT have all been shown to interact with pVHL in yeast, with the chaperonin TRiC/CCT potentially functioning as a ‘holdase’ for containing nascent pVHL until it can bind to its elongin cofactors (Feldman et al., 1999). Although TRiC/CCT hasn’t been directly implicated in degradation processes, some mutant forms of pVHL have also been shown to be able to refold after denaturing events in a chaperonin-independent way *in vitro* but not *in vivo* (Feldman et al., 2003; Yang et al., 2013).

In this study we present 20 novel pair swap mutations of pVHL that switch the position of two residues in the sequence and examine the protein products in yeast for a measure of their ability to persist in the cell. In particular, two mutations caused pVHL to be detected in much larger amounts than that of transfected wildtype pVHL: swapping the valine and glycine residues at positions 155 and 19 respectively and swapping the leucine and glutamic acid residues at positions 201 and 173. The first mutant directly affected a chaperone binding site, while the second did not mutate residues in any known chaperone or cofactor interaction motifs.

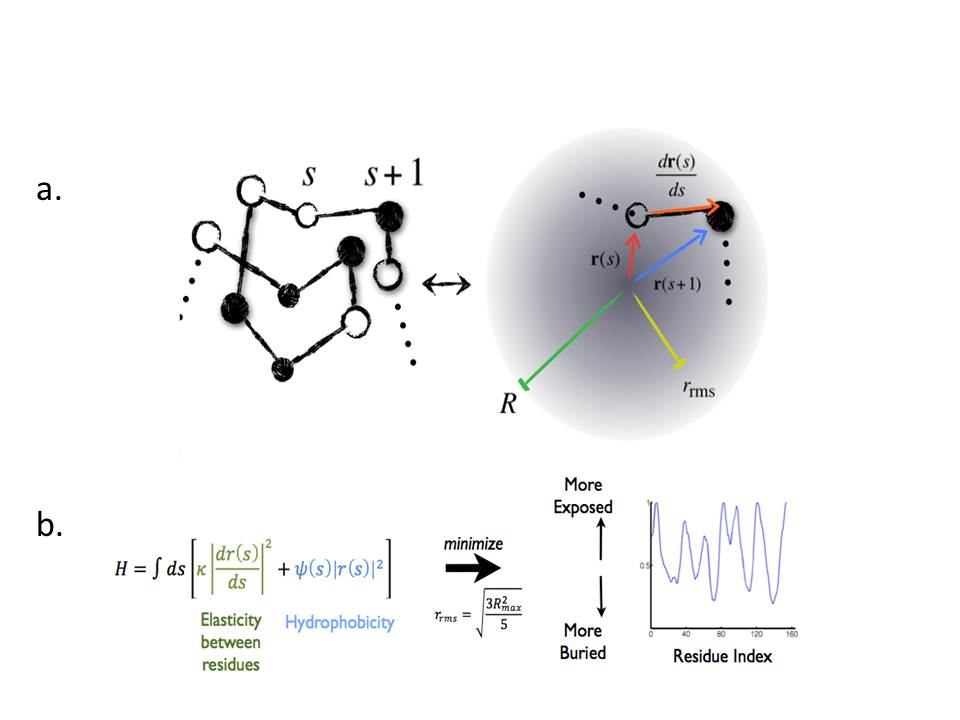
Initially, these results were surprising because both stabilizing mutations were chosen randomly as part of the control group, while the test group of ten mutants designed to exhibit more stability actually produced less-stable behavior on average. Different hypotheses attempt to explain how structurally abnormal proteins are recognized and marked for degradation, and half of the engineered mutants in this study were designed to have a low predicted structural variability compared to a control group chosen randomly. The underlying rationale behind choosing these mutants focused on predicting how easily a protein structure could access different structures at a low energy, with the idea that sequences displaying the least variability in predicted structure also might indicate increased thermodynamic and kinetic stability. The behavior of these mutants *in vivo* illustrates the difficulty in teasing apart the mechanisms for PQC substrate recognition. In the case of the ten designed mutants, the lower-than-expected observed amounts present in the cell may have been indicative that the mutated sequences were ‘kinetically trapped’ in one configuration that was more easily identified by the PQC system. In this case, our designed parameter to maintain protein structure despite small energetic fluctuations may have made these mutants more susceptible to degradation.

To explain the two significantly stable hits in the random control sample, the biophysical burial mode model was used to examine each of the twenty mutated sequences. These structural predictions were also compared to the wildtype sequence and its experimentally-determined structure. Strikingly, the burial mode model predicted that the mutated sequence L201-E173 would fold such that it would bury known hydrophobic chaperone interaction sites while exposing a motif known to be necessary and sufficient for binding to a homolog of its cofactor elongin C in yeast. The two chaperone interaction sites have been identified as binding to TRiC, with the first motif also similar to the known Hsp70 recognition motif of a short hydrophobic sequence flanked by charged residues. Furthermore, the cofactor elongin C exists in yeast despite the absence of other pVHL cofactors, and previous studies have indicated that it does not increase the half-life of wildtype pVHL. However, when this sequence was transfected into a yeast strain lacking the elongin homolog, levels of mutant pVHL were decreased to that of the wildtype protein, providing experimental evidence in tandem with computational observations.

This predicted conformation is similar to that of correctly-folded pVHL in humans, which could help explain why this mutant form is able to adopt a fold that is resistant to cellular degradation mechanisms. Furthermore, comparing this mutation to all possible mutated sequences in its well-folded region indicates that our experimental mutant is best able to follow this pattern of burying and exposing certain interaction motifs. These results indicate that hiding certain chaperone interaction sites and/or making cofactor interaction sites more surface-accessible while adopting an energetically-optimal fold could be potential mechanisms for pVHL to be protected from PQC machinery. Using pVHL as a model protein, the burial model has been shown to be capable of providing explanations of the relative likelihood of a mutated protein being degraded. Furthermore, this predictive capability was able to scan a large set of sequences in a relatively short time, providing rapid structural information about all possible mutations to narrow and refine experimental studies.

Both stabilization through modulation of chaperone interactions and through cooperative interactions with binding partners are implicated for pVHL in determining how the PQC recognizes and responds to misfolded variants.

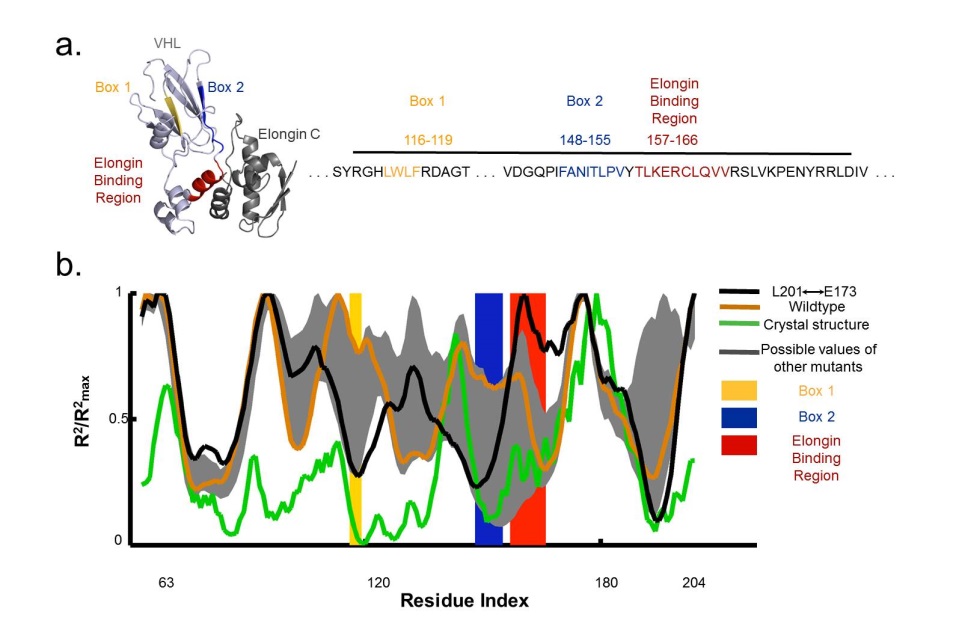
Expanding our computational and experimental analysis to other proteins besides pVHL may also prove informative in how the PQC responds to other marginally stable proteins. The well-characterized p53 protein is similar to pVHL in that it is small enough for burial trace analysis (393 amino acids), includes an N-terminal intrinsically disordered region, shows low thermodynamic and kinetic stability, and is characterized by a large number (>1000) of cancer-related mutations found in humans (Joerger and Fersht, 2007). Additionally, a recent study has indicated that p53 interacts with the same chaperonin, CCT, that is involved in pVHL folding (Trinidad et al., 2013). This protein’s similarity to pVHL, both in terms of its functional properties and its clinical importance, makes it an enticing target for burial/exposure analysis. Future directions include computationally predicting stabilizing mutations of this protein and experimentally testing them, based on initial promising results with the burial model and pVHL.



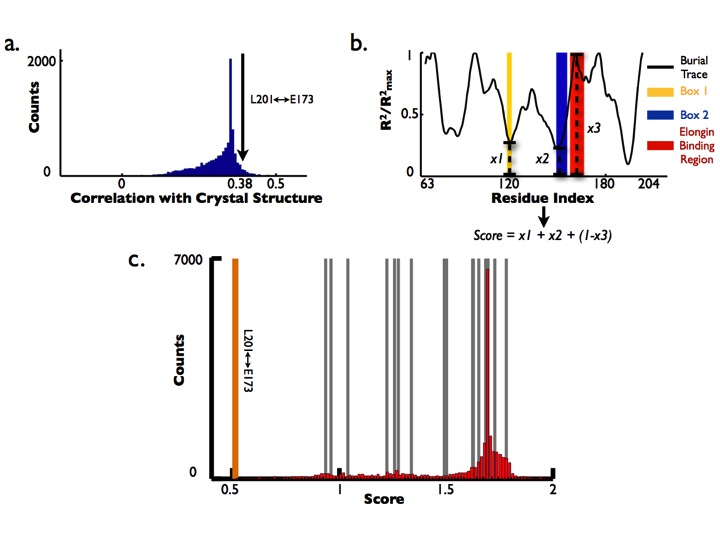
**Figure 1** (*a). An amino acid chain can be modeled as points connected by springs representing peptide bonds. The distance of the amino acid at position s from the center of mass of the protein is given by r(s), and the maximum distance R and the desired root mean square distance rrms are parameters to prevent steric clashing (b) Using the framework in (a), a Hamiltonian including energetic costs for stretching neighbor residues apart and putting hydrophobic residues on the outside of the protein can be minimized under our desired constraints to predict burial patterns.*

vhlFACSData992013.pdf

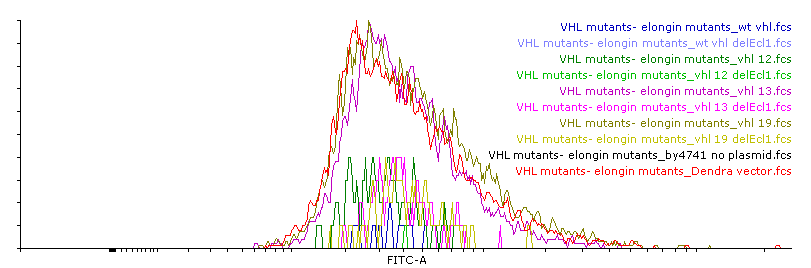
**Fig. 2.** *Fluorescence measurements by FACS analysis for all mutations except 16 and normalized to fluorescence of Dendra tag without attached pVHL. Fluorescence for the strain without a plasmid (by4741 no plasmid), the wildtype pVHL (VHL wt), and a known disease-causing mutation (vhl L158P) are also shown.*



**Figure 3** (*a)  The crystal structure of parts of pVHL can be obtained when the protein is bound in complex to cofactors like elongin C, which binds to VHL between residues 157-166.  VHL also is known to interact with chaperones in regions called Box 1 and Box 2. (b)  The L201/E173 mutation is predicted to bury its Box 1 and Box 2 regions, suggesting a conformation that is closer to the form the protein takes when stably bound as part of the VBC complex.  The elongin-interaction site of the mutation is also predicted to be highly exposed compared to both the wildtype and the other experimental mutations.*

****

**Figure 4** *(a)  Compared to all possible pair swap mutations in the region of VHL that is able to be crystallized (residues 63-204), L201/E173 is predicted to have a higher burial pattern correlation with the crystal structure than 93.82% of all other possible mutations. (b)  An illustration of how scores were generated for each possible mutant, where the smallest score would correspond to maximal burial of the Box 1 and 2 regions and maximal exposure of the elongin interaction site. (c)  Each possible pair swap mutation in the region [63,204] was scored according to Fig. 4b.  L201/E173 has the lowest (best) score out of ~10,000 mutations.  The locations of the other experimental mutants in the histogram are indicated by the gray lines, with L201/E173 marked by the orange line.*



**Fig. 5.** *Raw FACS histograms, indicating fluorescence on the x-axis and counts on the y-axis. The Dendra tag alone,wildtype pVHL, mutant 12 (E70-G19), mutant 13 (V155-G19), and mutant 19 (L201-E173) were expressed in both the presence and the absence of elongin C. The area under the histograms indicates overall fluorescence.*

**Literature Cited**

Amit, M., Weisberg, S.J., Nadler-Holly, M., McCormack, E.A., Feldmesser, E., Kaganovich, D., Willison, K.R., and Horovitz, A. (2010). Equivalent Mutations in the Eight Subunits of the Chaperonin CCT Produce Dramatically Different Cellular and Gene Expression Phenotypes. J. Mol. Biol. *401*, 532–543.

Botuyan, M.V., Mer, G., Yi, G.-S., Koth, C.M., Case, D.A., Edwards, A.M., Chazin, W.J., and Arrowsmith, C.H. (2001). Solution structure and dynamics of yeast elongin C in complex with a von hippel-lindau peptide. J. Mol. Biol. *312*, 177–186.

Chiti, F., and Dobson, C.M. (2006). Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem. *75*, 333–366.

Chiti, F., and Dobson, C.M. (2009). Amyloid formation by globular proteins under native conditions. Nat. Chem. Biol. *5*, 15–22.

England, J.L. (2011). Allostery in protein domains reflects a balance of steric and hydrophobic effects. Struct. Lond. Engl. 1993 *19*, 967–975.

Feldman, D.E., Thulasiraman, V., Ferreyra, R.G., and Frydman, J. (1999). Formation of the VHL-elongin BC tumor suppressor complex is mediated by the chaperonin TRiC. Mol. Cell *4*, 1051–1061.

Feldman, D.E., Spiess, C., Howard, D.E., and Frydman, J. (2003). Tumorigenic mutations in VHL disrupt folding in vivo by interfering with chaperonin binding. Mol. Cell *12*, 1213–1224.

Hansen, W.J., Ohh, M., Moslehi, J., Kondo, K., Kaelin, W.G., and Welch, W.J. (2002). Diverse Effects of Mutations in Exon II of the von Hippel-Lindau (VHL) Tumor Suppressor Gene on the Interaction of pVHL with the Cytosolic Chaperonin and pVHL-Dependent Ubiquitin Ligase Activity. Mol. Cell. Biol. *22*, 1947–1960.

Joerger, A.C., and Fersht, A.R. (2007). Structure–function–rescue: the diverse nature of common p53 cancer mutants. Oncogene *26*, 2226–2242.

Kaganovich, D., Kopito, R., and Frydman, J. (2008). Misfolded proteins partition between two distinct quality control compartments. Nature *454*, 1088–1095.

Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., and Ulrich Hartl, F. (2013). Molecular Chaperone Functions in Protein Folding and Proteostasis. Annu. Rev. Biochem. *82*, 323–355.

Knauth, K., Bex, C., Jemth, P., and Buchberger, A. (2006). Renal cell carcinoma risk in type 2 von Hippel–Lindau disease correlates with defects in pVHL stability and HIF-1α interactions. Oncogene *25*, 370–377.

Luby-Phelps, K. (2000). Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. Int. Rev. Cytol. *192*, 189–221.

Marcinowski, M., Rosam, M., Seitz, C., Elferich, J., Behnke, J., Bello, C., Feige, M.J., Becker, C.F.W., Antes, I., and Buchner, J. (2013). Conformational Selection in Substrate Recognition by Hsp70 Chaperones. J. Mol. Biol. *425*, 466–474.

McClellan, A.J., Scott, M.D., and Frydman, J. (2005). Folding and Quality Control of the VHL Tumor Suppressor Proceed through Distinct Chaperone Pathways. Cell *121*, 739–748.

Melville, M.W., McClellan, A.J., Meyer, A.S., Darveau, A., and Frydman, J. (2003). The Hsp70 and TRiC/CCT Chaperone Systems Cooperate In Vivo To Assemble the Von Hippel-Lindau Tumor Suppressor Complex. Mol. Cell. Biol. *23*, 3141–3151.

Nordstrom-O’Brien, M., van der Luijt, R.B., van Rooijen, E., van den Ouweland, A.M., Majoor-Krakauer, D.F., Lolkema, M.P., van Brussel, A., Voest, E.E., and Giles, R.H. (2010). Genetic analysis of von Hippel-Lindau disease. Hum. Mutat. *31*, 521–537.

Rosenbaum, J.C., Fredrickson, E.K., Oeser, M.L., Garrett-Engele, C.M., Locke, M.N., Richardson, L.A., Nelson, Z.W., Hetrick, E.D., Milac, T.I., Gottschling, D.E., et al. (2011). Disorder Targets Misorder in Nuclear Quality Control Degradation: A Disordered Ubiquitin Ligase Directly Recognizes Its Misfolded Substrates. Mol. Cell *41*, 93–106.

Rüdiger, S., Buchberger, A., and Bukau, B. (1997). Interaction of Hsp70 chaperones with substrates. Nat. Struct. Mol. Biol. *4*, 342–349.

Schoenfeld, A.R., Davidowitz, E.J., and Burk, R.D. (2000). Elongin BC complex prevents degradation of von Hippel-Lindau tumor suppressor gene products. Proc. Natl. Acad. Sci. *97*, 8507–8512.

Spiess, C., Miller, E.J., McClellan, A.J., and Frydman, J. (2006). Identification of the TRiC/CCT Substrate Binding Sites Uncovers the Function of Subunit Diversity in Eukaryotic Chaperonins. Mol. Cell *24*, 25–37.

Stebbins, C.E., Kaelin, W.G., Jr, and Pavletich, N.P. (1999). Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function. Science *284*, 455–461.

Sutovsky, H. (2004). The von Hippel-Lindau Tumor Suppressor Protein Is a Molten Globule under Native Conditions: IMPLICATIONS FOR ITS PHYSIOLOGICAL ACTIVITIES. J. Biol. Chem. *279*, 17190–17196.

Trinidad, A.G., Muller, P.A.J., Cuellar, J., Klejnot, M., Nobis, M., Valpuesta, J.M., and Vousden, K.H. (2013). Interaction of p53 with the CCT Complex Promotes Protein Folding and Wild-Type p53 Activity. Mol. Cell *50*, 805–817.

Yang, C., Huntoon, K., Ksendzovsky, A., Zhuang, Z., and Lonser, R.R. (2013). Proteostasis Modulators Prolong Missense VHL Protein Activity and Halt Tumor Progression. Cell Rep. *3*, 52–59.