Destabilization of the PAS domain core in the *Earlydoors* mutation of mouse *Period2* accelerates the circadian clock

Colby R. Sandate*

*Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064

Abstract

In mammalian systems, circadian timing is established by a series of interlinked feedback loops that involve the repressive proteins PERIOD (PER) and CRYPTOCHROME (CRY). PER:CRY complexes translocate into the nucleus to repress the circadian transcription factor CLOCK:BMAL1. After degradation of the PER and CRY proteins, the cycle begins anew. Regulation of PER stability through phosphorylation-dependent degradation plays a critical role in establishing the period of the mammalian clock. and can result in shortening or lengthening of the circadian period. Our collaborators performed a forward genetics screen in mice and identified a point mutation (I324N) located in the PAS domain core of PER2 that dramatically shortened the circadian period in mice. They named the mutant *Earlydoors* (*Edo*) after its shortened phenotype. Using a range of biophysical and biochemical techniques, we set out to determine the structural consequences the *Edo* mutation has for the PER2 PAS domain core. Our data demonstrate that the I324N *Edo* mutation in PER2 does not affect dimerization or the global conformation of the protein, but rather leads to lower thermodynamic stability of the protein, making it more susceptible to proteolysis. By integrating powerful genetics experiments and *in vivo* data from the Hastings and Nolan groups with a detailed biophysical study, we have now shown that changes in the intrinsic stability of PER2 can lead to severe circadian phenotypes.

Introduction

Overview of the Molecular Clock

Life evolved on a chaotic early Earth. In contrast to this initial chaos, the rhythmicity of the night and day cycle provided a pervasive source of stimuli. Most organisms adapted to this environmental cycle by developing molecular pathways that not only respond to the time of day, but also allows them to keep track of time in the absence of external stimuli. Called 'circadian rhythms', these pathways often control transcription of key gene products. Events in circadian regulation correspond to light/dark cycles, providing the organism with an intrinsic ability to keep track of time. Experimentally, circadian rhythms have been demonstrated to confer fitness to bacteria that contain them over other bacteria which do not, demonstrating a selective pressure to develop circadian rhythms.[1] The advantages conferred by circadian rhythms have also been demonstrated in higher organisms, such as plants.[2]

While circadian rhythms are present in most life forms, the molecular systems which govern them differ among organisms. The mammalian system is best described as a hierarchy of oscillators in which a transcriptional-translational feedback loop (TTFL), located within the neurons of the suprachaismatic nucleus (SCN), sits atop other oscillators located in somatic cells.[3,4] The SCN is a collection of neurons found within the anterior hypothalamus that acts as the master clock of the mammalian system to coordinate bodily rhythms through light stimuli from the retina (termed photoentrainment).[5] While other peripheral tissues contain their own circadian oscillators and can cycle independently, removal of the SCN oscillator results in a loss of synchrony within the molecular clocks of peripheral tissues.[6] For this reason, the SCN oscillator can be considered the 'master clock' of the organism, for which other clocks are aligned.

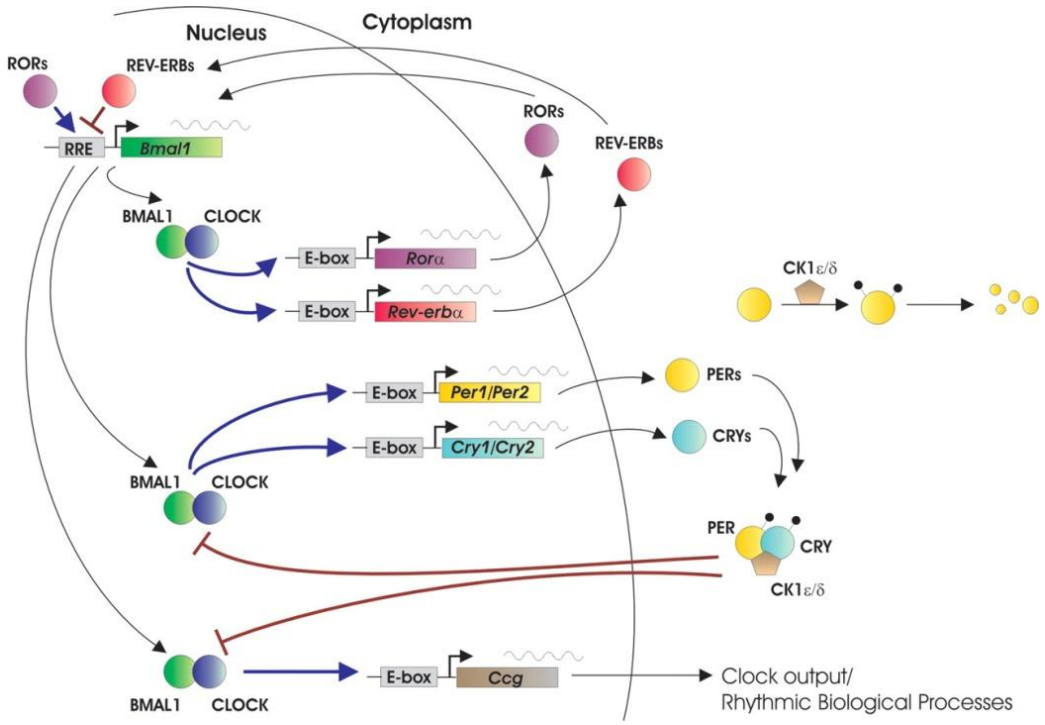


Figure 1. System diagram of the mammalian circadian clock. Mammalian circadian rhythms are generated by two feedback loops. The first is comprised of two transcription factors (CLOCK and BMAL1), which exert control over the expression of their own repressors (PERs and CRYs). An important regulatory event that partially determines the length of the circadian period, involves the phosphorylation of PERs by CK1 ε/δ, functionalizing them for dimerization with CRYs, as well as for degradation. A secondary feedback loop involves nuclear receptors (REV-ERBs and RORs) which control the expression of BMAL1.[3]

The mammalian clock is comprised of positive and negative elements (Figure 1). The positive elements consist of two members of the basic helix-loop-helix (bHLH) PER-ARNT-SIM (PAS) domain transcription factor family: CLOCK (Circadian Locomotor Output Cycles Kaput) and BMAL1 (Brain and Muscle ARNT-like 1).[7] These two proteins dimerize within the cytoplasm, translocate into the nucleus[8] and bind to E-box elements of DNA, activating transcription of clock-controlled genes. Several of those clock-controlled genes include the Period proteins (PER1, PER2) and two Cryptochrome proteins (CRY1 and CRY2), which comprise the negative elements of the oscillator. PER and CRY are phosphorylated prior to nuclear translocation, where they repress their own transcription by interacting with CLOCK:BMAL1.[9] The ~24 hour periodicity of this oscillator is governed by posttranslational modifications of the repressors such as phosphorylation-mediated nuclear-entry, and ubiquitin-initiated degradation of the proteins

themselves. PER2 and CRY1 are phosphorylated by Casein Kinase 1ε and 1δ (CK1 ε and CK1 δ) prior to nuclear entry[10], and the E3 ligase β-TCRP 1/2 mark PER2 for degradation.[11]

PAS Domains

While a system-wide perspective is necessary to provide context to our findings, the work presented in this thesis is primarily focused on protein structure. Domain architecture plays a central role in protein structure, as it forms a stabile scaffold for protein functionality. Three of the four core clock proteins contain PAS domains: CLOCK, BMAL1, and PER2. PAS domains were first discovered in a bioinformatics-based study of protein sequence conservation, resulting in the identification of a minimal sequence motif in primary sequence and tertiary space (Figure 2).[12] The PAS-A and PAS-B domains of CLOCK and BMAL1 share many of the components present in the minimal conserved structure shown in Figure 2, but differ in a few respects. PAS-A domains usually have an additional helix at their N-terminus, commonly referred to as the A'α helix.[13] This A'α helix docks onto the β-sheet of the PAS domain core. PAS-B domains are usually smaller and more condensed. Some PAS-B domains contain a C-terminal helix, called the Jα helix, which can dock onto the β-sheet of α-helix side of the PAS-B domain. This allows for regulation of PAS domain function.

PAS domain-containing proteins typically behave as a type of sensor: PAS input domains will detect some kind of physical or chemical stimuli and then regulate the activity of output domains, such as catalysis or DNA binding (as is the case in CLOCK and BMAL1).[14] This kind of regulation is found in many interesting biological processes; including nitrogen fixation in *Rhizobia*, photoropism in plants, and ion channel gating in vertebrates. As mentioned previously, most of the core circadian proteins in mammals contain tandem (i.e. side-by-side) PAS domains.

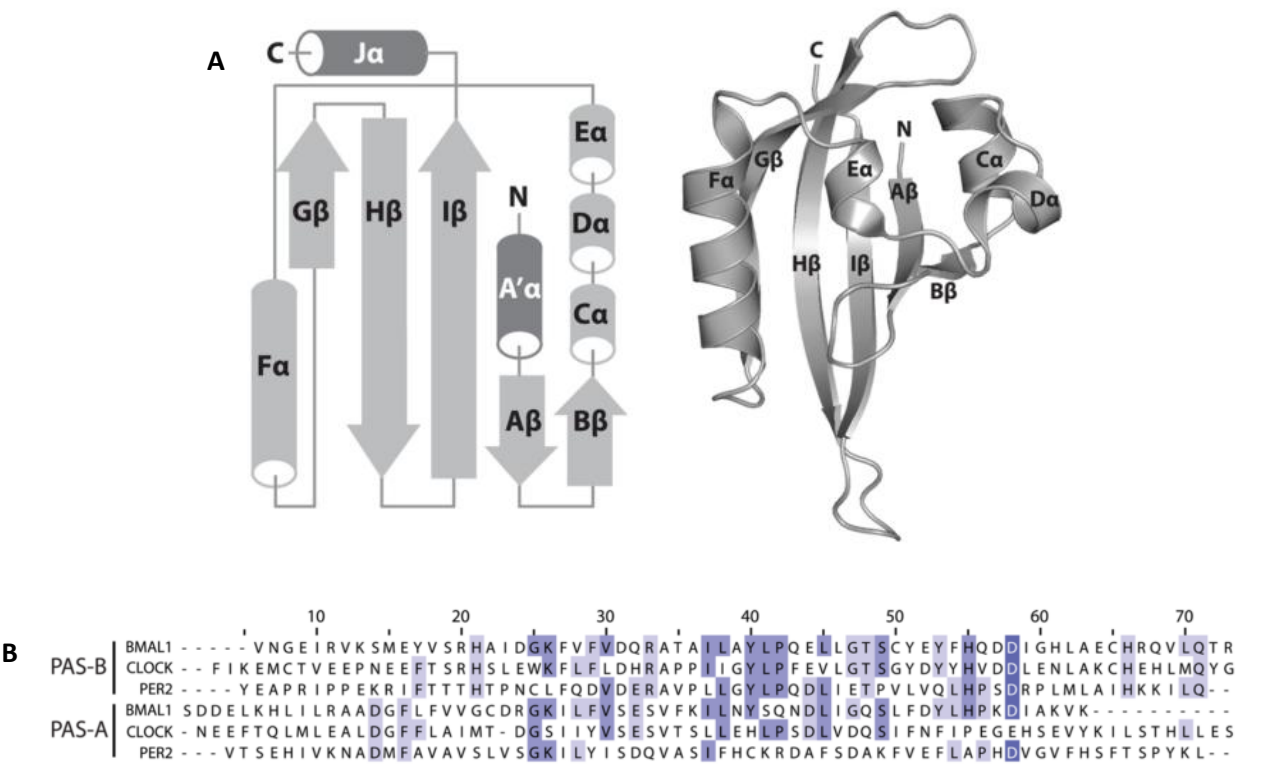


Figure 2. The minimally conserved PAS domain. A) Left: 2D schematic of secondary structural components of the minimal conserved PAS domain (light grey), including secondary structure components most commonly associated with PAS-A (A'α helix) and PAS-B (Jα helix) domains (dark grey). Barrels: α helix; arrow: β-strand. Secondary structure components are labeled alphabetically, beginning at the N-terminus. Right: 3D cartoon visualization of the minimal conserved PAS domain structure from BMAL1 PAS-B. B) Sequence conservation in residues located in PAS-A and PAS-B domains between the proteins BMAL1, CLOCK and PER2.[11]

PERIOD: Repressor and Key Player

PERs comprise half of the PER:CRY repressive complex. Deletion of PER1, PER2 or CRY1 destroys functionality of the negative arm of the molecular clock, leading to arrythmicity.[15,16] While it was originally thought that CRY proteins are the components responsible for allowing nuclear translocation of the repressive complex, it was shown in 2001 that it is actually PER proteins that are the rate-limiting species.[17] In mammals, PER proteins come in three variants: PER1, PER2 and PER3. Experiments conducted *in vitro* and *in vivo* show that PER3 cannot maintain cycling in the absence of PER1 and PER2 and is associated with output functions[18,19]. The remainder of the discussion on PERs focus on PER2, as the mutant *Edo* PER2 is the focus of this study.

In 2009, a high-resolution crystal structure of a mPER2 homodimer was published (the 'm' in front of PER2 distinguishes the molecule as being from murine origin), providing unprecedented detail of the molecular arrangement of the PAS domain core (Figure 3)[20]. The crystal structure shows that after a short region of disordered N-terminal residues, PER2 contains a PAS-A domain which is connected to a corresponding PAS-B domain through a highly conserved linker segment. The PAS-A domain does not interact appreciably with the PAS-B domains of either molecule. This lack of contacts allows the PAS-A domain to be flexible, giving it the ability to interact with other proteins.[11] For example, the PAS-A domain contains a LXXLL (where 'L' depicts leucines residues and 'X' can be any amino acid) recognition sequence that allows PER2 to interact with several nuclear receptors.[21] The C-terminus of PER2 is mainly disordered and contains a CK1 ε/δ binding site and a CRY1/2 site, which allow for posttranslational modification of PER2 and PER:CRY dimerization.

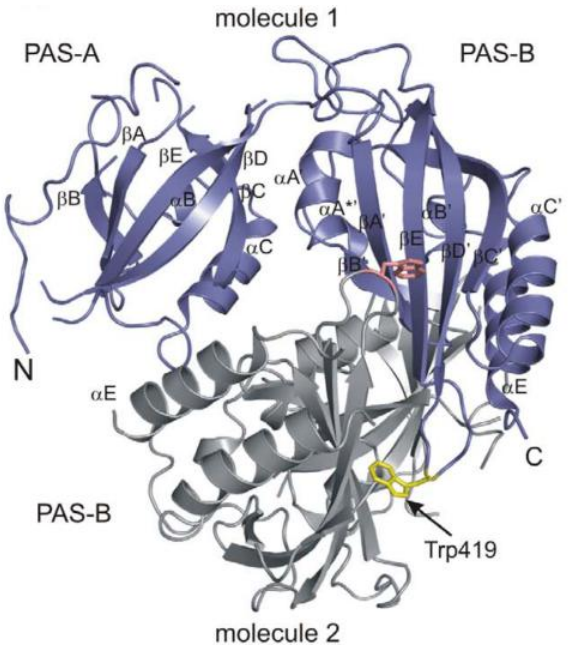


Figure 3. The structure of PER2. A) Crystal Structure of PER2 PAS domain core homodimer. Molecule 1 (blue) shows PAS-A and PAS-B domain as well as highly conserved Trp419 (yellow)[20]. Molecule 2 is shown in grey.

True to its name, the PER2 protein seems to have a direct effect on determining the period of the TTFL. Since PER2 is subjected to posttranslational modifications which affect its positioning – in both time and space within the TTFL – mutations that influence the ease at which PER2 is modified can directly lead to interesting circadian phenotypes. For example: a missense mutation in hPER2, the human analogue of mouse PER2, was implicated in familial advanced sleep phase syndrome (FASPS), a genetic disorder characterized by a shortened circadian period and an earlier onset of sleeping behaviors.[22] The missense mutation caused a serine to glycine substitution in the critical CK1 ε region of hPER2, which diminished the rate of phosphorylation of hPER2 and decreased the length of the circadian period. Conversely, the circadian period can also be lengthened through chemical modulation of PER2 posttranslational modification. Treatment of cycling cells with kinase inhibitors has been shown to increase the circadian period anywhere from 24 to 48 hours, demonstrating the staggering spectrum of plasticity of PER regulation by kinases.[23]

Studying PER2 presents a powerful opportunity for circadian period modulation. Additionally, the identification of mutations in PER2 that affect circadian period is an important step in mechanistically describing how circadian timing is established. Such an opportunity arose when Michael H. Hastings and Patrick M. Nolan, our collaborators at the University of Cambridge, performed a forward genetics screen in mice using 1-ethyl-1-nitrosourea (ENU) mutagenesis to search for mutant mice with altered circadian phenotypes. They quantified the circadian periods of the mice by monitoring daily wheel-running activity. Using this method, they identified a point mutation in PER2 that dramatically shortened the circadian period (Figure 4A). They named the mutant *Earlydoors* (*Edo*) after its shortened phenotype and mapped the mutation to an isoleucine-asparagine substitution (I324N) in the conserved linker region between the two PAS domains of PER2 (Figure 4B). Through breeding experiments, they established that the *Edo* mutation is semi-dominant and that heterozygotes displayed a period of 23 hours while the *Edo* homozygote displayed a period of 22 hours (Figure 4C). Interestingly, when HEK293 cells co-expressed with both CK1 ε and *Edo* PER2, and then were subsequently treated with the translation inhibitor cycloheximide (CHX), a time-course western blot revealed that *Edo* PER2 degraded faster than the wild type (WT) protein (Figure 4D). While the Hastings and Nolan groups were able to show that *Edo* PER2 degrades faster than the WT *in vivo*, they were unable to assign a mechanistic explanation for why this particular mutation results in decreased stability.

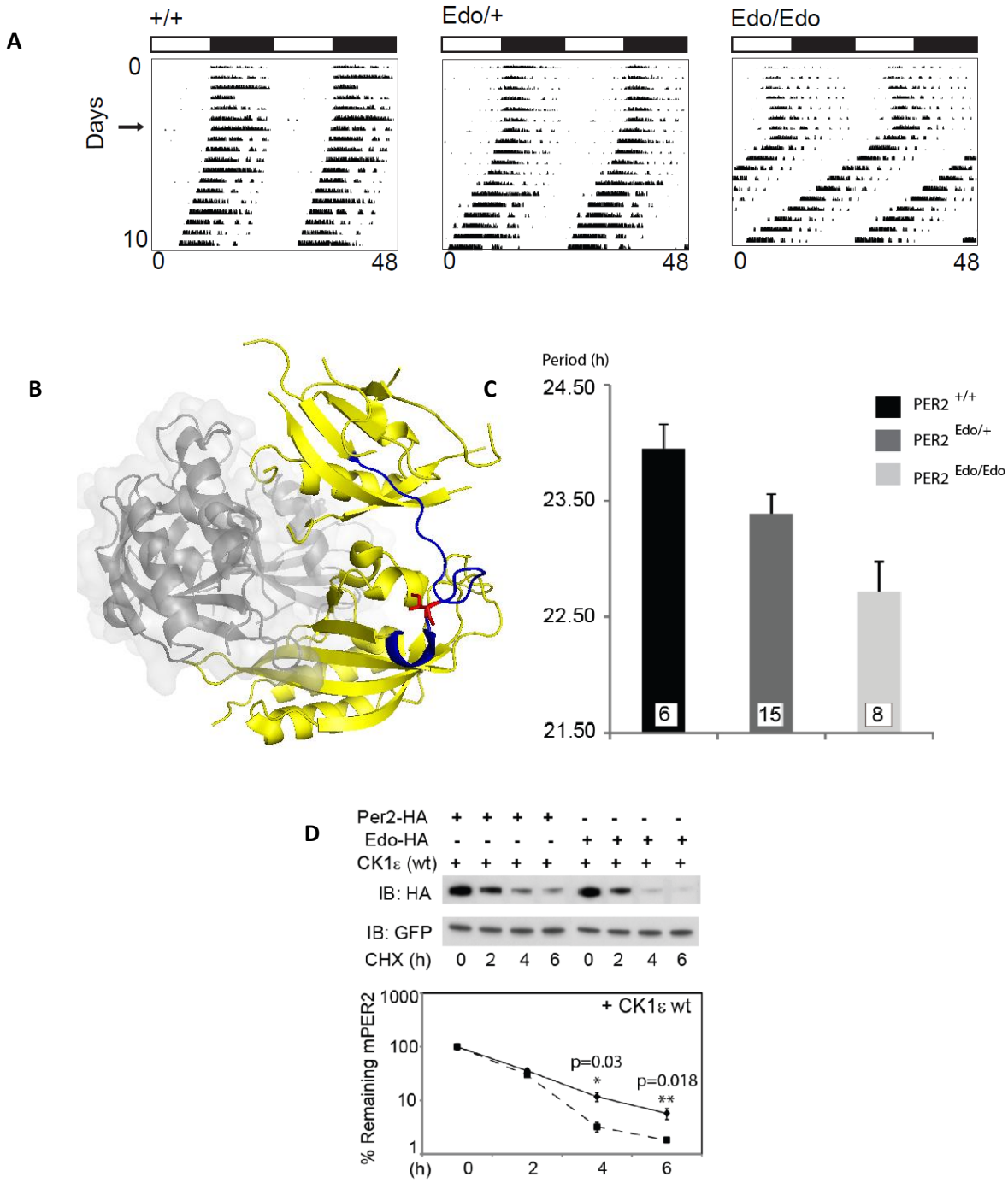


Figure 4. Data obtained from the Hastings and Nolan groups. A) Wheel running activity for heterozygous dominant mice (left), heterozygous mice (middle) and homozygous *Edo* mice (right). Mice hetero or homozygous for *Edo* displayed a shortening of activity period which was dose dependent on number of copies of *Edo* gene received. Activity is depicted as black tick marks on each row (representing one day) under light conditions (white) and dark conditions (black). Data are double plotted to better visualize circadian rhythms. B) Crystal structure (3GDI) of PER2 homodimer showing site of I324N mutation (red), within the interdomian linker (blue). Molecule 1 appears in yellow while molecule 2 appears in black. C) Histogram displaying circadian period recorded for WT and Edo mice with error and number of mice sampled (white box). D) (Top) Time course western blot showing that Edo PER2 degrades much faster than WT. (Bottom) Quantification of western blot data.

Experimental Methods

Protein Expression and Purification

A construct of the PAS domain core of *Edo* PER2 (residues 170-473 which comprise the PAS-A and PAS-B domain) including a TEV-cleavable N-terminal His_6 tag, was used to transform the *Escherichia coli* (*E. coli*) Rosetta(DE3) strain. Bacterial cultures were induced to express PER2 *Edo* by the addition of 0.5 mM isopropyl β-D-1-thiogalctopyranoside (IPTG) once OD_{600} readings reached ~0.8. After addition of IPTG, bacterial growths were incubated for an additional 16 hours at 18 °C while shaking. Cultures were then centrifuged at 4000 RPM for 30 min at 4 °C to collect bacteria. Pellets were resuspended in buffer containing 50 mM Tris, 300 mM NaCl and 20 mM imidazole (pH 7.5) before undergoing cell lysis via cell extruder. Soluble proteins were purified from the cell lysates though a Ni^{2+}-nitrilotriacetic acid (Ni-NTA) affinity chromatography column (purchased from QIAGEN), which was also selective for the His_6 tagged protein. The His_6 tag was cleaved from the *Edo* PER2 protein through addition of His_6–TEV protease and incubation at room temperature for 4-6 hours. TEV cleavage was followed by an additional run over a Ni-NTA affinity column to remove cleaved tag and protease. Final purification was accomplished through size-exclusion chromatography on a Superdex 75 prep grade column previously equilibrated in a buffer containing 25 nM HEPES pH 7.5, 200 mM NaCL, 5 mM DTT. Protein solutions were then concentrated down to 10 mg/ml using an Amicon spin concentrator equipped with a 10 kDa molecular weight cutoff filter.

SAXS Data Acquisition and Analysis

Proteins were freshly purified through size-exclusion chromatography and prepared in a buffer solution containing 25mM HEPES pH 7.5, 200 mM NaCl and 5 mM DTT and 2% glycerol for protection against radiation damage. Protein samples were then diluted to three concentrations (1, 3, and 5 mg/ml). Data was collected on the SIBYLS beamline at the Advanced Light Source, Lawrence Berkeley National Laboratory. Samples were irradiated at

different time intervals to check for concentration dependence of scattering and radiation damage. A buffer blank was also irradiated, allowing for buffer subtraction from the protein signal. Data sets containing minimal concentration dependence and radiation damage were merged to create a more complete dataset (greater coverage of q values, the momentum transfer which is a function of scattering angle and incident X-ray wavelength) for further analysis. Experimental data were compared to a theoretical intensity dataset derived from the PER2 crystal structure (PDB: 3GDI) using the online FoXS server[24]. Descriptive values (radius of gyration, Porod volume, etc.) were determined using ATSAS suite of programs.[25] P(r) functions and appropriate D_{max} values were generated using the GNOM program through ATSAS. These P(r) functions were used to calculate ten independent *ab initio* solution envelopes using the DAMMIN program and were then aligned and averaged using DAMAVER, another ATSAS set of programs. Final averaged and aligned solution envelopes were visualized using UCSF's Chimera software[26] by docking a crystal structure of PER2 into the envelope. The PDB file containing the crystal structure was modified using the online AllosMod-FoXS server[27,28] to model in regions of the structure lacking electron density due to disordered loops and highly flexible regions.

Limited Proteolyis

Protein solutions of 1.5 mg/ml suspended in buffer containing 25 mM HEPES pH 7.5, 200 mM NaCL and 5 mM DTT underwent 1 hour proteolysis reactions using sequencing-grade proteases chymotrypsin and trypsin (Promega). Separate reactions contained the following mass (w/w) ratios: chymotrypsin (1:25, 1:50, 1:100) or trypsin (1:100, 1:250, 1:500). Reaction quenching was accomplished through the addition of 2X SDS Laemmli buffer (Bio-Rad), resulting in a 1:1 dilution. Samples were then boiled at 95°C for 5 minutes before loading onto a 18% SDS-PAGE gel. Gel samples were ran for 50 minutes at 200 Volts before visualization by Coomassie stain.

Differential Scanning Fluorimetry

The thermodynamic stability of protein folding was assessed using a miniaturized thermal denaturation assay[29] utilizing the hydrophobic SYPRO Orange dye. A 10 μM protein solution containing 25 mM HEPES pH 7.5, 200 mM NaCl, 5 mM DTT and 2X SYPRO Orange (Sigma Aldrich) was added (in six replicates), to a 96-well PCR plate with a final reaction volume of 50 μL. The PCR plate was sealed using an optically clear film (Invitrogen), and then inserted into a ViiA7 Real-Time PCR instrument and ran using an altered Melt Curve program[27]. Fluorescence data were analyzed in Excel, to trim data at high and low temps far away from the sigmoidal region and, before determining melt temperatures through non-linear regression of Boltzmann sigmoidal denaturation curves using Graph Pad Prism 6. The data was fitted to the following equation:

$$Fluorescence = Bottom + \frac{(Top - Bottom)}{1 + \exp\left(\frac{T_m - temp}{slope}\right)}$$

Where 'Top' is the fluorescence value when the protein is completely unfolded, and 'Bottom' is the fluorescence value when the protein is folded.

Results and Discussion

The *Edo* Mutation Does Not Affect Dimerization

We began our investigation by analyzing the mutant protein's ability to form homodimers, as this is an established function of PAS domains.[12] Tryptophan 419 (W419) is essential for formation of PER2 homo/heterodimers, as mutagenesis studies abolishing W419's aromatic properties have resulted in loss of dimer formation.[20] As such, W419 and surrounding residues in the PAS-B domain make up the dimerization interface of PER2 and make it highly likely that any perturbations in PER2 dimerization would have to affect this region. The site of

the mutation, I324N, is 95 residues away in primary sequence space. However, tertiary space is much more informative of any possible interactions of I324N to the dimerization interface. Analysis of the crystal structure reveals that the site of the mutation is nearly 36 Å away from W419 (Figure 5A), making the possibility of any distant interaction highly unlikely. From this information, we predicted that dimerization would not be affected in *Edo* PER2.

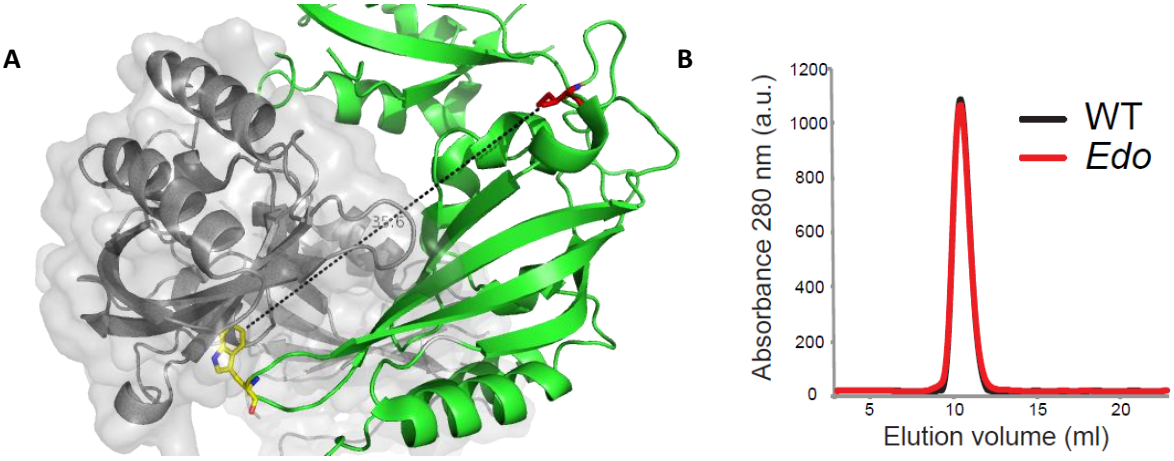


Figure 5. I324N does not affect dimerization. A) Distance measurement between I324 (red) and W419 (yellow). Molecule 1 appears in green while molecule 2 appears in black. B) Absorbance$_{280}$ vs elution volume plot from size-exlusion chromatography.

Size-exclusion chromatography suggests that *Edo* PER2 forms a homodimer in solution. Hennig *et al* had previously shown, through their own size-exclusion chromatography, that PER2 travels through SEC as a homodimer in solution.[20] *Edo* PER2 and WT both elute at nearly the same volume when ran on a Superdex 75 analytical column (Figure 5B), which indicates that *Edo* PER2 is also traveling as a homodimer. Aggregate proteins would elute earlier in size exclusion, and Monomeric molecules would elute much later. Additionally, the calculated molecular weight (MW) prediction by small angle X-ray scattering (SAXS) was 66 kDa for WT and 67 kDa for *Edo* PER2, strongly suggesting that both proteins existed in their homodimeric form while in solution (Table 1), as their values were very similar. For comparison, the MW for a single monomer of PER2 is ~35 kDa while the dimer is ~70 kDa. The deviation of the calculated MW from SAXS from the actual MW of the dimer can be explained by its

dependence on I(0) (see below). Since I(0) is dependent on concentration, the calculated MW will be affected variances in protein concentration.

SAXS Analysis and Lack of Evidence for Large-Scale Structural Change

After finding that the *Edo* mutation does not affect PER2 dimerization, we began searching for evidence of large-scale structural change in the protein fold. I324N is located in a hydrophobic pocket within the interdomain linker between the PAS-A and PAS-B domains. The I324N mutation could possibly destabilize contacts between the two domains, as an isoleucine (a hydrophobic residue) to asparagine (a polar residue) substitution could produce a drastic change in local environment (Figure 6). A decoupling between the PAS-A and PAS-B domains resulting from the *Edo* mutant would dramatically change the shape of the protein in solution, yet still allow for dimerization, as the dimerization interface is primarily located on the PAS-B domain and does not involve any contacts in the PAS-A domain.

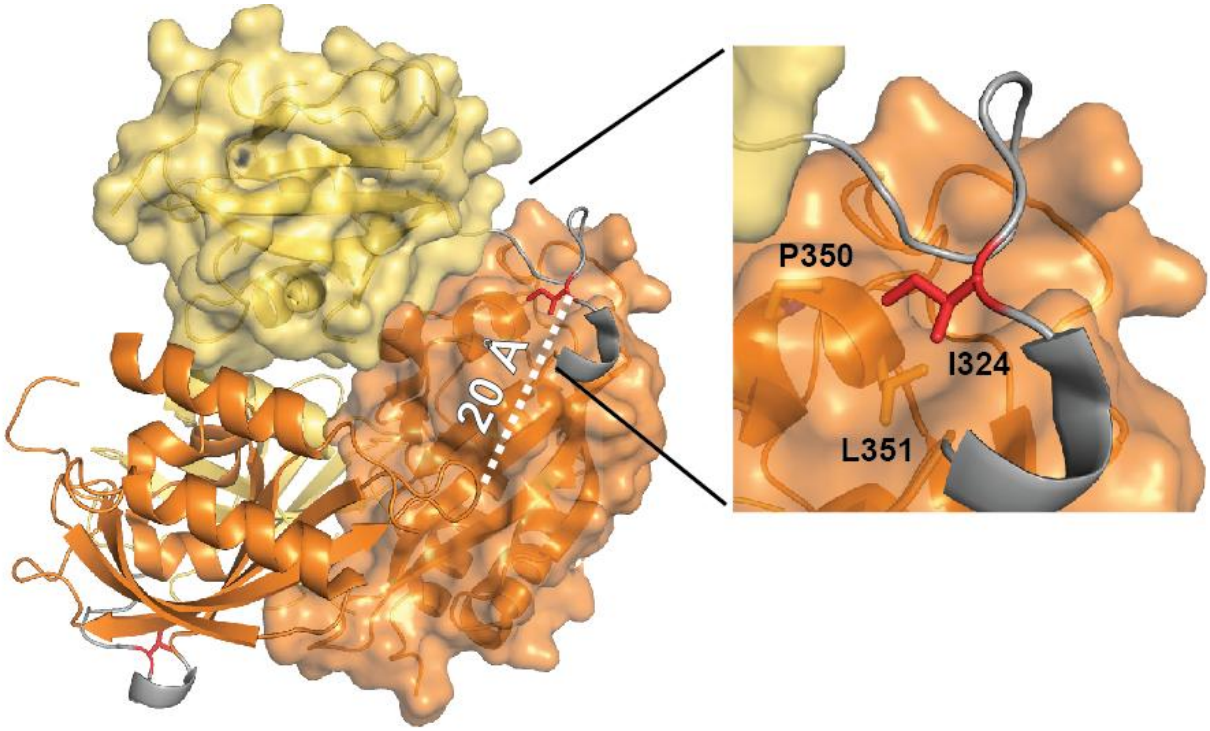


Figure 6. Close up of I324N mutation in hydrophobic pocket. The site of the *Edo* mutation (red) is shown in proximity to nearby hydrophobic residues (P350, L351). Molecule 1 is depicted as a surface representation, while molecule 2 is depicted as a cartoon. PAS-A domains (yellow), PAS-B domains (orange) and interdomian linker (grey) are also shown in this figure. Adapted with permission from Professor Carrie Partch.

We used SAXS to search for evidence of global conformational change. SAXS is a similar method to X-ray diffraction in that both use a collimated X-ray beam to irradiate samples (Figure 7). However, SAXS only collects scattering (not diffraction) data from small angles, which limits the structural information it can collect to low resolutions.[30] Despite this limitation, SAXS offers several key advantages over X-ray diffraction. SAXS does not require biological samples to be packed into single crystals but can easily resolve samples in solution. This allows for a much easier sample preparation, as the researcher does not have to search for a set of conditions that gives diffraction quality crystals. While SAXS can provide solution structures of dilute protein samples, it is easily biased by particle interactions as well as radiation damage. Before any conclusions can be drawn from SAXS analysis, the data must be meticulously checked for quality control, as discussed below.

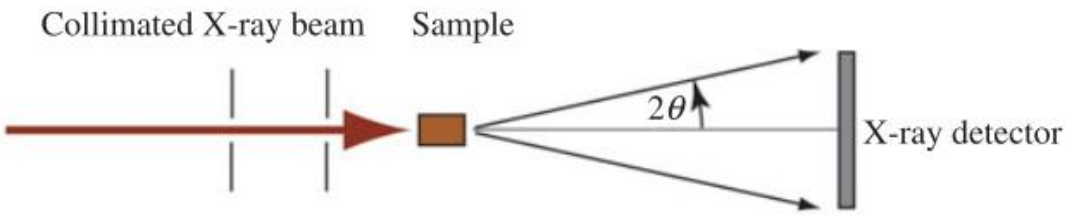


Figure 7. Simplified schematic of SAXS experiment. A collimated X-ray beam (red), generated from a Synchrotron source, irradiates a protein solution sample (orange). Only scattering data from small angles (2θ) is collected by the X-ray detector for analysis.

Scattering intensity curves for both WT and *Edo* PER2 are nearly identical (Figure 8A). Both curves contain multiple features, seen as peaks and valleys, which indicate that these were well-behaved samples. Proteins that are aggregating non-uniformly would give a scattering curve with a featureless exponential decay. Additionally, curves generated from different concentrations of both protein solutions were superimposable, meaning that different scattering particles were not interacting with each other in solution.

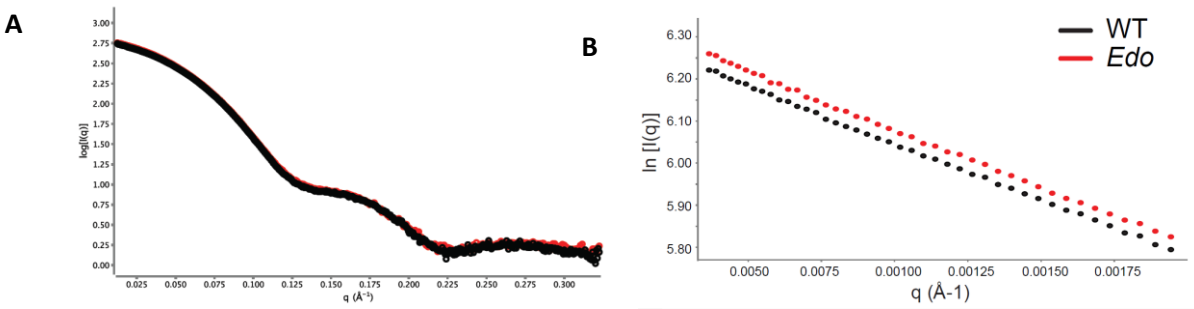


Figure 8. SAXS intensity data and Guinier approximation for *Edo* PER2 and WT. A) Superimposed SAXS intensity curve for *Edo* PER2 (red) and WT (black). B) Guinier approximation plot in small q region.

	I(0) (Å)	R_g (Å)	D_{max} (Å)	MW (Da)	Porod Volume (Å^3)
Edo	535.1±0.2	27.000±0.004	71	6.69E+04	1.11E+05
WT	517.5±0.2	26.900±0.004	71	6.59E+04	1.07E+05

Table 1. Descriptive values obtained from SAXS data analysis.

Another important quality control measure is the Guinier approximation, which assesses data quality for aggregation, polydispersity, and success of background subtraction by analyzing the linearity of small q intensity values (defined as q x R_g < 1.3). In general, a linear Guinier approximation plot with a negative slope is desirable. The data sets for WT and *Edo* PER2 are highly linear in this region, giving us confidence to move forward in our analysis. From the Guinier approximation, we extrapolated I(0), the predicted intensity value at q = 0, as well at the radius of gyration (R_g), the square root of the average squared distance of each scatter from the particle center. *Edo* PER2 gave an I(0) value of 535.1 ± 0.2 while WT returned a value of 517.5 ± 0.2 (Table1). The differences between these two values could be explained by small differences in protein concentration between the two solutions, as scattering intensity is partially dependent on concentration. Both proteins had highly similar R_g values with *Edo* PER2 having 27 Å and WT having an R_g value of 26.9 Å. The Porod volume, the calculated volume excluded from the solvent by a single scattering particle, was 1.11x10^5 Å^3 for PER2 *Edo* and 1.07x10^5 Å^3 for WT.

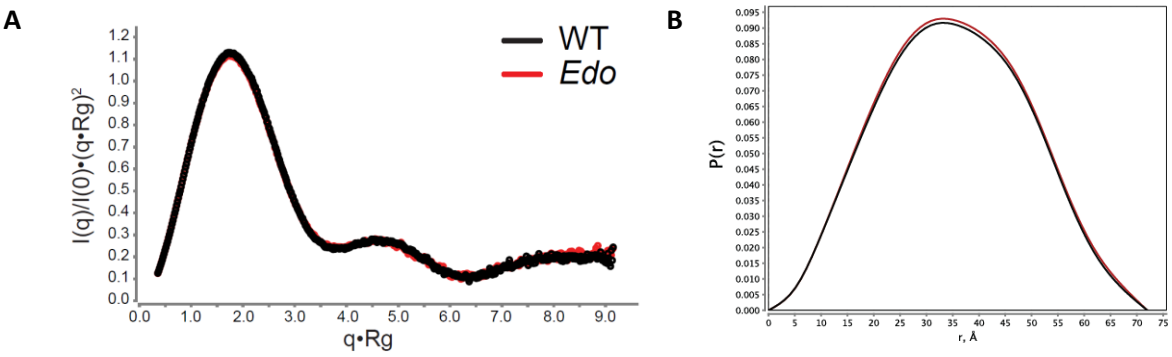


Figure 9. Kratky plot and P(r) distribution function. A) Superimposed Kratky plot. B) Superimposed refined P(r) function generated from GNOM for *Edo* PER2 (red) and WT (black).

The Kratky plot provides some structural information, and is therefore another indicator of data quality. A bell-shaped curve in their Kratky plots is indicative of a globular protein. If proteins are well folded, their curves should move toward lower intensity values at larger q•Rg values. Such is the case with both *Edo* PER2 and WT (Figure 9A). While the Kratky plot provides some useful information, most of the structural insights from SAXS begin with the creation of P(r) distribution function (Figure 9B). *Edo* PER2 and WT both generate a P(r) function that contains only a single peak – which is typical of relatively spherical globular proteins. Proteins with multiple discrete domains tend to generate P(r) functions with multiple peaks, representing distance restraints within each of the discrete domains. Although PER2 contains two tandem PAS domains, the tight packing of the PAS domain core into a heterodimer likely prevents each domain from being resolved individually, resulting in a single peak.

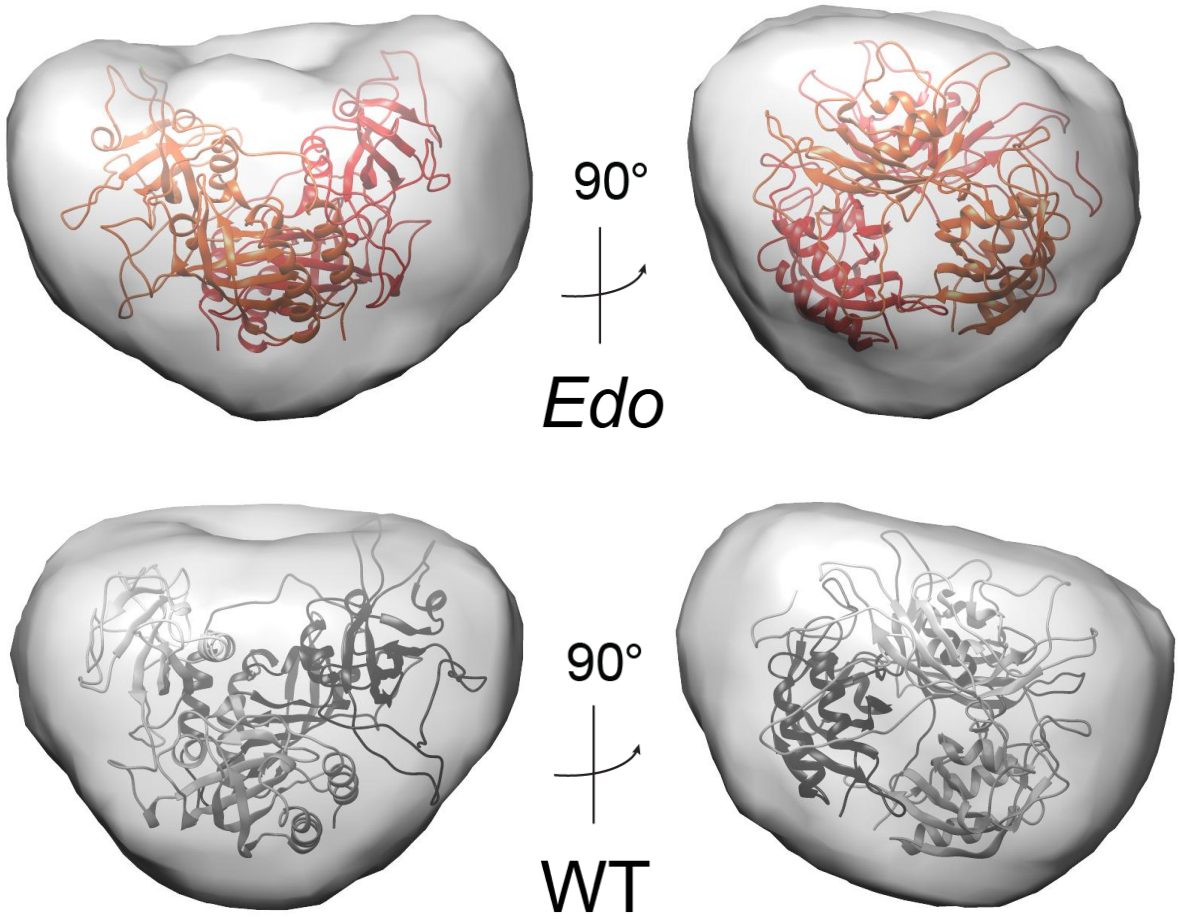


Figure 10. Final SAXS solution envelopes. Envelopes were generated from P(r) function with crystal structure 3GDI docked for both *Edo* PER2 (red) and WT (grey).

From the P(r) function, we generated solution envelopes that provide low-resolution structural information about the overall shape of *Edo* and WT PER2. The final merged envelopes, averaged from 10 independent calculations, are shown above (Figure 10). The SAXS solution envelopes model the surfaces of the average protein conformation sampled when irradiated. Because of conformational averaging, and the tight packing brought about by crystallization, some volume of the envelope will not be filled by the docked crystal structure. From visual inspection, it is clear that while the *Edo* PER2 envelope shows slightly more

features that the WT envelope, there is little difference between the two structures. Both are shaped like an oblate sphere and both readily accept a modified (See Appendix A) 3GDI crystal structure of a PER2 homodimer for docking. This result, when combined with the similar values obtained for R_g and D_{max} between the two proteins, does not support our hypothesis that the *Edo* mutation results in global conformational change within the PAS domain core of PER2. While these data refute the possibility of global change between the two PAS domain, it does not preclude small conformational change occurring at the site of the mutation – the interdomain linker.

Edo Mutation Affects PER2 Stability

Following the logic that the I324N mutation may lead to small-scale perturbations near the interdomain linker, we performed limited proteolysis to probe changes in solvent accessibility between the *Edo* PER2 and WT peptide chains. Limited proteolysis is a technique that readily recognizes areas of increased flexibility or solvent exposure in proteins. It relies on the partial digestion of proteins by limiting quantities of proteases, assayed for discrete time points through quenching of the reaction after short time intervals. Because of the limited protease concentrations and time periods available for the reaction, only areas that are solvent exposed and accessible to the protease are cleaved from the larger fragment. We used both chymotrypsin and trypsin to account for different protease specificity. Of the two, only trypsin showed a difference between *Edo* PER2 and WT (See Appendix B). The results of the trypsin experiment are shown below (Figure 11A). With increasing concentrations of trypsin, a preferential accumulation of an ~18 kDa band is seen in the *Edo* PER2 sample. While this band does occur in WT samples, it appears at higher concentrations of protease instead of accumulating much earlier in *Edo* PER2 samples. Additionally, a band slightly below the 35 kDa non-digested protein band appears only in WT samples. Together, these data suggest that the flexibility of the PAS domain core is altered by the *Edo* mutation.

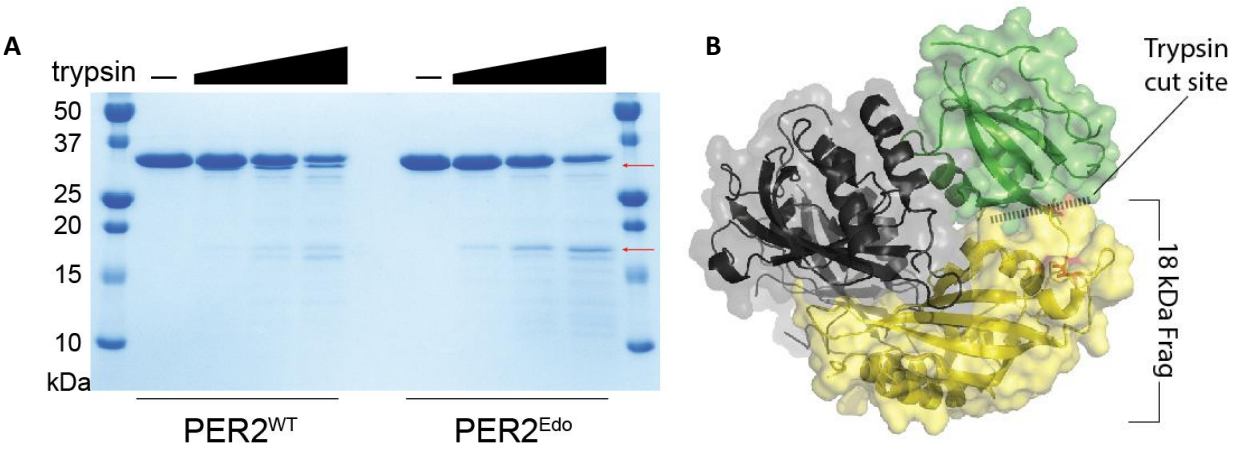


Figure 11. Limited trypsin proteolysis data. A) Products of trypsin limited proteolysis visualized on an 18% SDS-PAGE gel. Red arrows indicate distinct species observed between WT and *Edo* PER2. Wedges indicate increasing concentration of trypsin used for a 60 minute digestion at room temperature. B) Schematic of PER2 homodimer depicting trypsin cut site (dashed line) as well as proposed identity of the 18 kDA fragment (yellow) that was cleaved from molecule 1 (green) preferentially in Edo PER2. Molecule 2 appears in black.

Once trypsin proteolysis revealed a difference between *Edo* PER2 and WT, we searched the primary sequence of PER2 for trypsin cut sites that corresponded to the fragment sizes we observed. Our search identified a trypsin cut site within the interdomian linker and near I324N (See Appendix C)**,** that would produce an 18 kDa fragment comprising the entire PAS-B domain and most of the interdomain linker (Figure 11B). A second cut site is located in a C-terminal loop that connects the Jα helix to the adjacent PAS-B domain. We interpret these findings as suggesting that the I324N mutation perturbs the interdomain linker from remaining buried within the hydrophobic pocket between the tandem PAS domains. Instead, the hydrophilic asparagine residue in the *Edo* mutant likely causes much of the linker backbone to become solvent exposed and much more available for proteolysis compared to WT. The C-terminal loop connecting the Jα helix to the remaining PAS-B domain is probably one of the most solvent exposed trypsin cut sites in WT protein, but is probably less exposed compared to the perturbed interdomain linker in *Edo* PER2. Preferential cleavage of the cut site located near I324N in *Edo* PER2 over the C-terminal loop is consistent with the absence of the < 35 kDa band in *Edo* PER2 samples.

If the I324N mutation in *Edo* PER2 leads to perturbation of the interdomain linker docking into the hydrophobic pocket on the PAS-B domain, it could follow that the entire fold of *Edo* PER2 is destabilized compared to WT. To test this hypothesis, we used differential scanning fluorimetry to measure the melting temperature (T_m) of the two purified proteins (Figure 12). Melting temperature is a measure of thermodynamic stability; molecules that melt at a higher temperature are more stable. SYPRO Orange, the fluorescent dye used in this experiment, gives minimal signal in aqueous buffers but exhibits a dramatic increase in quantum yield in hydrophobic environments. SYPRO Orange readily binds to hydrophobic regions in proteins, when they are accessible. Since proteins unfold and exhibit their hydrophobic cores when exposed to increasing temperature, the fluorescence emitted from SYPRO Orange can be used as proxy for protein unfolding. Using a non-linear fitting of the data to the Boltzman sigmoidal equation, this experiment revealed a difference in T_m values between WT and *Edo* PER2 (54.6 ± 0.4 °C and 49.4 ± 0.3 °C, respectively), resulting in an overall ΔΔG of 0.66 kcal/mol. The lower T_m for *Edo* PER2 strongly suggests that the I324N mutation destabilizes the *Edo* PER2 fold. This agrees with our previous finding that the interdomain linker is less buried within the hydrophobic pocket between the tandem PAS domains in *Edo* PER2.

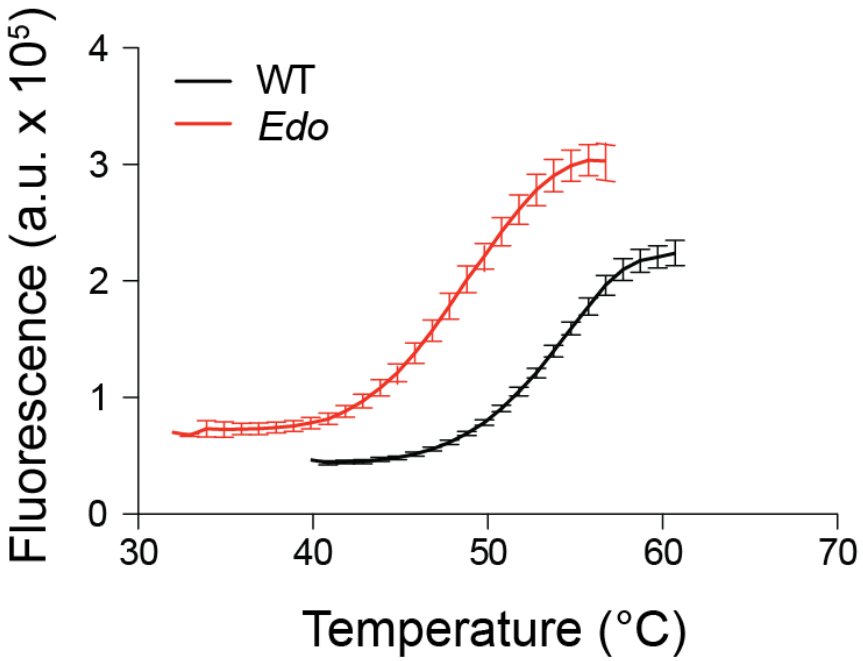


Figure 12. Non-linear Boltzmann sigmoidal curve fit of differential scanning flourimetry data. The mean fluorescence data (n=6) is shown here as non-linear fitted lines with SD error bars. T_m values occur at the midpoint of the sigmoidal curve.

Conclusion

Our results show that the I324N mutation in *Edo* PER2 does not affect dimerization or cause a large-scale conformational change in the protein fold. Instead, the shortening of the circadian period in *Edo* mice is likely due to the combined effects of destabilization of the PAS domains caused by the loss of a hydrophobic residue. This allows the more solvent-exposed linker region to become more susceptible to proteolysis, causing *Edo* PER2 to degrade much faster than WT. The latter is in agreement with the *in vivo* data from the Hastings and Nolan groups showing that *Edo* PER2 degrades much faster that WT in living cells (Figure 4D). It is remarkable that such a minor change in flexibility, brought upon by a single amino acid substitution, can lead to a significant change in circadian period. These data underline the importance of PER2 stability in establishing circadian timing, in line with previous studies.[31,32]

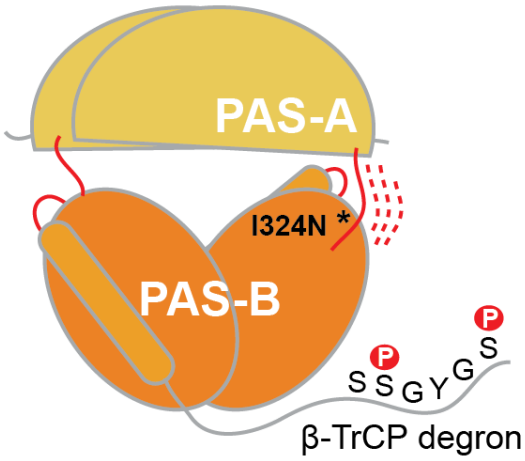


Figure 13. Proposed model for increased flexibility of interdomian linker in *Edo* PER2. Perturbation of interdomain linker (dashed red lines), leads to a less stable protein fold, increased susceptibility to proteolysis, and increased accessibility to the β–TrCP degron sequence.

While destabilizing the intrinsic stability of the *Edo* PER2 protein fold will most likely lead to a shorter protein half-life *in vivo*, the increased flexibility of the interdomian linker may have additional consequences for regulated protein degradation. A β-TrCP degron sequence is located immediately downstream of the PAS-B domain, which serves as the primary site by which the E3 ligase β-TrCP is recruited to PER2 to control its stability (Figure 13)[33,34]. By increasing the flexibility of the interdomain linker, the β-TrCP degron sequence may become more available and might allow *Edo* PER2 to be marked for degradation more rapidly *in vivo*. While this question was not investigated in this study, it is a promising target for future research.

References

1. Ouyang, Y., Andersson, C. R., Kondo, T., Golden, S. S. & Hirschie, C. Resonating circadian clocks enhance fitness in cyanobacteria. **95,** 8660–8664 (2009).

2. Dodd, A. N. *et al.* Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* **309,** 630–633 (2005).

3. Ko, C. H. & Takahashi, J. S. Molecular components of the mammalian circadian clock. *Hum. Mol. Genet.* **15,** 271–277 (2006).

4. Hastings, M. H., Reddy, A. B. & Maywood, E. S. A clockwork web: circadian timing in brain and periphery, in health and disease. *Nat. Rev. Neurosci.* **4,** 649–661 (2003).

5. Yamazaki, S. *et al.* Resetting central and peripheral circadian oscillators in transgenic rats. *Science* **288,** 682–685 (2000).

6. Yoo, S.-H. *et al.* PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 5339–5346 (2004).

7. Reppert, S. M. & Weaver, D. R. Coordination of circadian timing in mammals. *Nature* **418,** 935–941 (2002).

8. Kondratov, R. V. *et al.* BMAL1-dependent circadian oscillation of nuclear CLOCK: Posttranslational events induced by dimerization of transcriptional activators of the mammalian clock system. *Genes Dev.* **17,** 1921–1932 (2003).

9. Kume, K. *et al.* mCRY1 and mCRY2 are essential components of the negative limb of the circadian clock feedback loop. *Cell* **98,** 193–205 (1999).

10. Akashi, M., Tsuchiya, Y., Yoshino, T. & Nishida, E. Control of intracellular dynamics of mammalian period proteins by casein kinase I ε (CKI ε) and CKI δ in cultured cells. *Mol. Cell. Biol.* **22,** 1693–1703 (2002).

11. Stojkovic, K., Wing, S. S. & Cermakian, N. A central role for ubiquitination within a circadian clock protein modification code. *Front. Mol. Neurosci.* **7,** 1–7 (2014).

12. Huang, Z. J., Edery, I. & Rosbash, M. PAS is a dimerization domain common to Drosophila period and several transcription factors. *Nature* **364,** 259–262 (1993).

13. Gustafson, C. L. & Partch, C. L. Emerging models for the molecular basis of mammalian circadian timing. *Biochemistry* **54,** 134–149 (2015).

14. Möglich, A., Ayers, R. a. & Moffat, K. Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure* **17,** 1282–1294 (2009).

15. Bae, K. *et al.* Differential functions of mPer1, mPer2, and mPer3 in the SCN circadian clock. **30,** 525–536 (2001).

16. Liu, A. C. *et al.* Intercellular coupling confers robustness against mutations in the SCN circadian clock network. *Cell* **129,** 605–616 (2007).

17. Lee, C., Etchegaray, J. P., Cagampang, F. R., Loudon, a S. & Reppert, S. M. Posttranslational mechanisms regulate the mammalian circadian clock. *Cell* **107,** 855–867 (2001).

18. Akashi, M. *et al.* A positive role for PERIOD in mammalian circadian gene expression. *Cell Rep.* **7,** 1056–1064 (2014).

19. Kucera, N. *et al.* Unwinding the differences of the mammalian PERIOD clock proteins from crystal structure to cellular function. *Proc. Natl. Acad. Sci.* **109,** 3311–3316 (2012).

20. Hennig, S. *et al.* Structural and functional analyses of PAS domain interactions of the clock proteins Drosophila PERIOD and mouse period2. *PLoS Biol.* **7,** 0836–0853 (2009).

21. Schmutz, I., Ripperger, J. a., Baeriswyl-Aebischer, S. & Albrecht, U. The mammalian clock component PERIOD2 coordinates circadian output by interaction with nuclear receptors. *Genes Dev.* **24,** 345–357 (2010).

22. Toh, K. L. *et al.* An hPer2 phosphorylation site mutation in familial advanced sleep phase syndrome. *Science* **291,** 1040–1043 (2001).

23. Isojima, Y. *et al.* CKIepsilon/delta-dependent phosphorylation is a temperature-insensitive, period-determining process in the mammalian circadian clock. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 15744–15749 (2009).

24. Schneidman-Duhovny, D., Hammel, M., Tainer, J. a. & Sali, A. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.* **105,** 962–974 (2013).

25. Konarev, P. V., Petoukhov, M. V., Volkov, V. V. & Svergun, D. I. ATSAS 2.1, a program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **39,** 277–286 (2006).

26. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).

27. Weinkam, P., Pons, J. & Sali, a. Structure-based model of allostery predicts coupling between distant sites. *Proc. Natl. Acad. Sci.* **109,** 4875–4880 (2012).

28. Schneidman-Duhovny, D., Hammel, M. & Sali, A. FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res.* **38,** 540–544 (2010).

29. Eswar, N. *et al.* Analysis of protein stability and ligand interactions by thermal shift assay. *Curr. Protoc. Protein Sci.* **Chapter 2,** Unit 2.9 (2001).

30. Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. a. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular

structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40,** 191–285 (2007).

31. Maywood, E. S. *et al.* Tuning the period of the mammalian circadian clock: additive and independent effects of CK1εTau and Fbxl3Afh mutations on mouse circadian behavior and molecular pacemaking. *J. Neurosci.* **31,** 1539–1544 (2011).

32. Meng, Q. J. *et al.* Setting clock speed in mammals: the CK1ε tau mutation in mice accelerates circadian pacemakers by selectively destabilizing PERIOD proteins. *Neuron* **58,** 78–88 (2008).

33. Eide, E. J. *et al.* Control of mammalian circadian rhythm by CKIepsilon-regulated proteasome-mediated PER2 degradation. *Mol. Cell. Biol.* **25,** 2795–2807 (2005).

34. Ohsaki, K. *et al.* The role of β-TrCP1 and β-TrCP2 in circadian rhythm generation by mediating degradation of clock protein PER2. *J. Biochem.* **144,** 609–618 (2008).

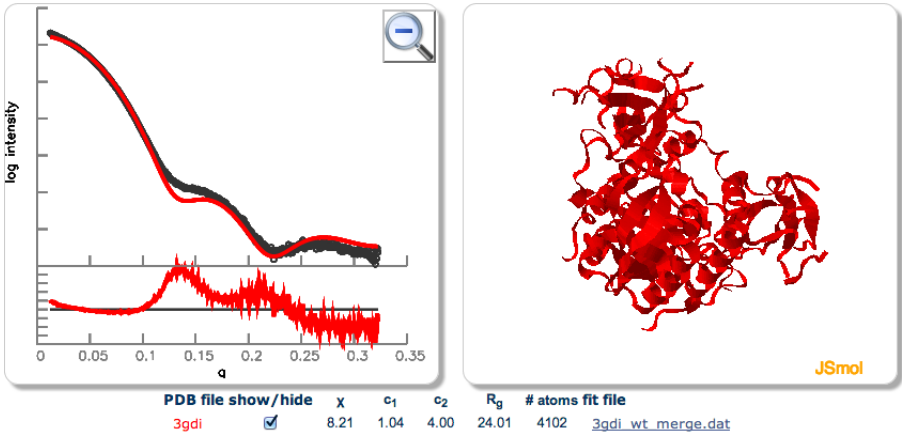
Appendices

Appendix A

Modification of 3GDI crystal structure

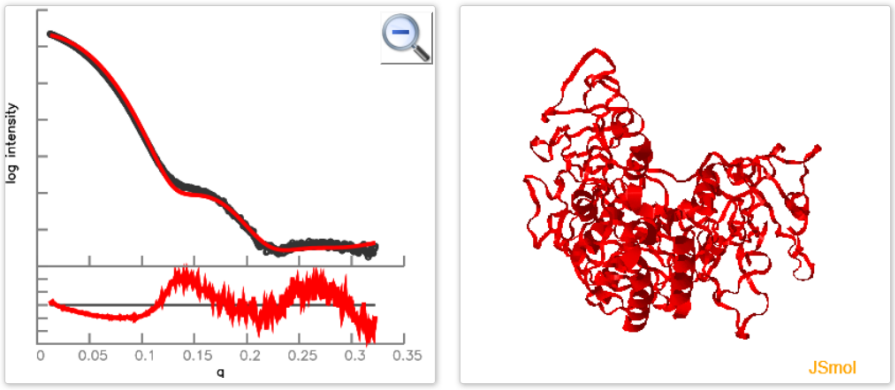
The crystal structure of PER2 (PDB: 3GDI), was modified using the online server AllosMod-FoXS, to improve fitting to experimental SAXS data. 3GDI without modifications (A: right) is missing several flexible loops in the crystal structure. This gives a poor χ-value of 8.21(A: left), where χ is a statistical measure of fit. After loops had been added in (B:right), the χ-value drops to 0.11 and appears to fit the experimental data quite well (B: left). Although the χ-value for 3GDI with loops is less than 1, which is usually an indication of overfitting, any reduction in χ-values much greater than one is considered an improvement.

A



PDB file	show/hide	χ	c_1	c_2	R_g	# atoms	fit file
3gdi	☑	8.21	1.04	4.00	24.01	4102	3gdi_wt_merge.dat

B

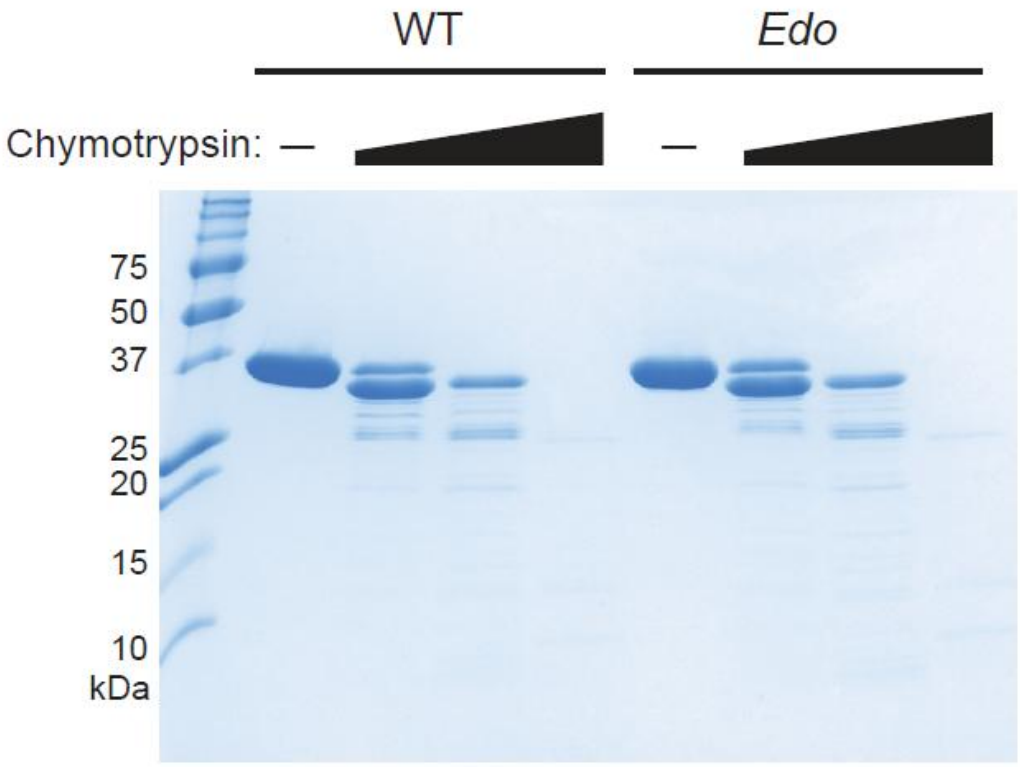


PDB file	show/hide	χ	c_1	c_2	R_g	# atoms	fit file
per2loops	☑	0.11	1.04	-0.49	25.82	4946	per2loops_3gdi_mutmerge.dat

Appendix B

Limited chymotrypsin proteolysis data

Products of chymotrypsin limited proteolysis visualized on an 18% SDS-PAGE gel. Wedges indicate increasing concentration of trypsin used for a 60 minute digestion at room temperature. No significant difference was observed between *Edo* PER2 and WT.



Appendix C

Amino acid sequence of Edo PER2 with trypsin cut sites

Sequence is shown, from left to right, from N-terminus to C-terminus. Trypsin cut sites are highlighted in blue and I324N is highlighted in red. Residues cooresponding to the PAS-A domain are shown in yellow, those cooresponding to the PAS-B domain are shown in organge, and those found within the interdomain linker are shown in dark grey. Residues remaining from N-terminal His_6 tag are shown in light grey.

GAMDPSYSMEQVEGITSEYIVKNADMFAVAVSLVSGKILYISNQVASIFHCKKDAFSDAKFVEFLA
PHDVSVFHSYTTPYKLPPWSVCSGLDSFTQECMEEKSFFCRVSVGKHHENEIRYQPFRMTPYLVKV
QEQQGAESQLCCLLLAERVHSGYEAPRNPPEKRIFTTTHTPNCLFQAVDERAVPLLGYLPQDLIETP
VLVQLHPSDRPLMLAIHKKILQAGGQPFDYSPIRFRTRNGEYITLDTSWSSFINPWSRKISFIIGRHK
VRVGPLNEDVFAAPPCPEEKTPHPSVQELTEQIHRLLMQPVP