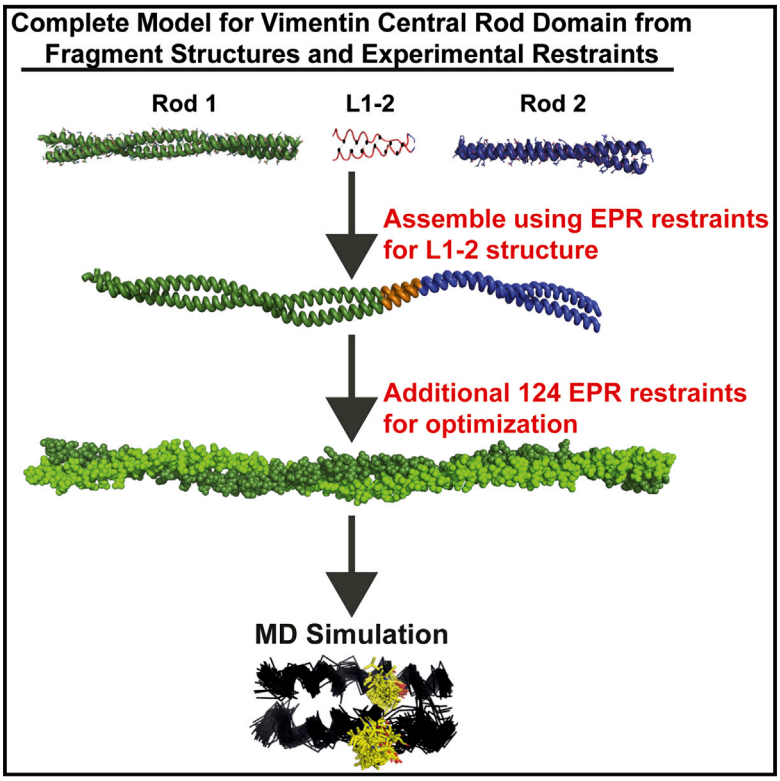
Article

Structure

# Completion of the Vimentin Rod Domain Structure Using Experimental Restraints: A New Tool for Exploring Intermediate Filament Assembly and Mutations

## Graphical Abstract




## Highlights

- EPR-based restraints for backbone structure and optimizing models of protein assemblies
- L1-2 is mostly helical with the two chains forming a heterogeneous structure
- L1-2 through rod 2A contains distinct levels of order and lacks supercoiling
- Assembled model provides basis for exploring IF dynamics and assembly

## Authors

David D. Gae, Madhu S. Budamagunta, John F. Hess, Robert M. McCarrick, Gary A. Lorigan, Paul G. FitzGerald, John C. Voss

## Correspondence

pgfitzgerald@ucdavis.edu (P.G.F.), jcvoss@ucdavis.edu (J.C.V.)

## In Brief

The linker 1–2 region of the intermediate filament protein has thus far eluded structural determination. Gae et al. use a combination of EPR spectroscopy and molecular modeling to describe linker 1–2 and unite the separate elements of the vimentin rod domain into a single experimentally based model.

Gae et al., 2019, Structure 27, 1–14
October 1, 2019 
https://doi.org/10.1016/j.str.2019.07.011

CellPress

Please cite this article in press as: Gae et al., Completion of the Vimentin Rod Domain Structure Using Experimental Restraints: A New Tool for Exploring Intermediate Filament Assembly and Mutations, Structure (2019), https://doi.org/10.1016/j.str.2019.07.011

Structure
Article



# Completion of the Vimentin Rod Domain Structure Using Experimental Restraints: A New Tool for Exploring Intermediate Filament Assembly and Mutations

David D. Gae,[1] Madhu S. Budamagunta,[2] John F. Hess,[3] Robert M. McCarrick,[4] Gary A. Lorigan,[4] Paul G. FitzGerald,[3,*] and John C. Voss[2,5,*]

[1]Department of Surgery, School of Medicine, University of California San Francisco, San Francisco, CA 94118, USA
[2]Department of Biochemistry & Molecular Medicine, School of Medicine, University of California Davis, Davis, CA 95616, USA
[3]Department of Cell Biology and Human Anatomy, School of Medicine, University of California Davis, Davis, CA 95616, USA
[4]Department of Chemistry and Biochemistry, Miami University, Oxford, OH 45056, USA
[5]Lead Contact
*Correspondence: pgfitzgerald@ucdavis.edu (P.G.F.), jcvoss@ucdavis.edu (J.C.V.)
https://doi.org/10.1016/j.str.2019.07.011

## SUMMARY

**Electron paramagnetic resonance (EPR) spectroscopy of full-length vimentin and X-ray crystallography of vimentin peptides has provided concordant structural data for nearly the entire central rod domain of the protein. In this report, we use a combination of EPR spectroscopy and molecular modeling to determine the structure and dynamics of the missing region and unite the separate elements into a single structure. Validation of the linker 1–2 (L1-2) modeling approach is demonstrated by the close correlation between EPR and X-ray data in the previously solved regions. Importantly, molecular dynamic (MD) simulation of the constructed model agrees with spin label motion as determined by EPR. Furthermore, MD simulation shows L1-2 heterogeneity, with a concerted switching of states among the dimer chains. These data provide the first ever experimentally driven model of a complete intermediate filament rod domain, providing research tools for further modeling and assembly studies.**

## INTRODUCTION

The intermediate filament (IF) protein family is a large and diverse group of cytoskeletal proteins found in the cytoplasm and nucleus that have been linked to a multitude of inheritable disorders (Omary, 2009). Although no complete IF protein has been crystallized, analysis of common sequence motif across the IF family have generated a model of cytoplasmic IF proteins that is structurally distinguished by a large central rod domain dominated by a coiled-coil structure but interrupted by short "linker" regions that lack a clear structural motif (reviewed in Herrmann et al., 2007). The central rod domain is flanked by amino terminal "head" and C-terminal "tail" domains (Parry and Steinert, 1999; Steinert et al., 1984, 1985). Within the central rod domain, analysis identified a common pattern of coiled-coil domains, rods 1A, 1B, 2A, and 2B, separated by short linkers predicted not to be coiled coil (Hanukoglu and Fuchs, 1983; Steinert et al., 1984). With a strategy termed "Divide and Conquer," Strelkov, Herrmann, Aebi and coworkers solved the crystal structure of several peptides derived from human vimentin (Chernyatina and Strelkov, 2012; Herrmann et al., 2000; Nicolet et al., 2010; Strelkov et al., 2001, 2002). Strelkov and coworkers showed a coiled-coil structure over a large part of vimentin rod 2B, confirming long-standing predictions (Strelkov et al., 2002). Other crystal structures have revealed surprises, such as a structure of rod 1A with a single $\alpha$ helix, rather than coiled-coil helices (Strelkov et al., 2002). Similarly, the peptide sequence of rod 2A through linker 2 (L2) adopted a surprising structure: a pair of parallel helices intertwined into a tetrameric antiparallel alignment (Nicolet et al., 2010). Separately, each pair of parallel helices supported the hypothesis made by Parry (2006) that rod 2A and L2 adopt a parallel helices structure in intact IFs, and were consistent with data from site-directed spin labeling combined with electron paramagnetic resonance spectroscopy (SDSL-EPR) published at the same time showing straight parallel chains for this region (Hess et al., 2006).

Using human vimentin as a model for IF structure, SDSL-EPR (Aziz et al., 2012; Hess et al., 2002, 2004, 2006) and X-ray crystallography (Aziz et al., 2012; Chernyatina and Strelkov, 2012; Herrmann et al., 2000; Nicolet et al., 2010; Pang et al., 2018; Strelkov et al., 2001, 2002) have been highly complementary (Chernyatina et al., 2016), confirming a coiled-coil structure for rod 1B, linker 2, and rod 2B. The only remaining uncharacterized region of the central rod domain is termed linker 1–2 (L1-2), located near the middle of the central rod domain. Analysis of amino acid sequence homology across all IF classes led Conway and Parry (1988) to conclude that L1-2 contained a partially conserved structure that was possibly $\beta$ strand. A $\beta$ strand fold for the L1-2 region has also been predicted by modeling (Guzenko and Strelkov, 2018; Qin et al., 2009). It has been further proposed that L1-2 is more dynamic than other domains, assuming an $\alpha$-helical conformation within assembled filaments, with a propensity for $\beta$-like structure in unassembled subunits (Parry and Smith, 2010).

In this study, we use a combination of SDSL-EPR and a novel model-building process to analyze the L1-2 region of vimentin,

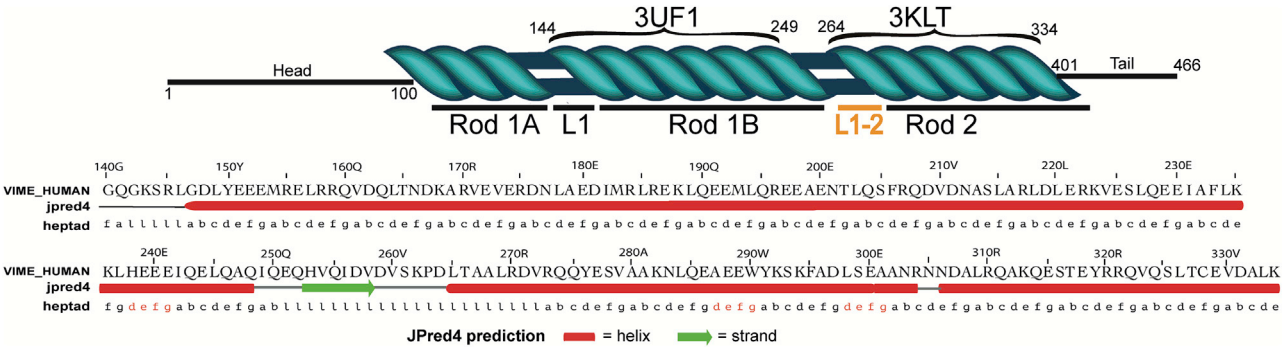



**Figure 1. Schematic Diagram of Vimentin Sequence, Secondary Structural Prediction, and EPR/X-Ray Structural Analysis**

The top diagram provides an overview of the structural regions within the rod domain, along with the fragments that have been defined by X-ray crystallography. The orange color highlights the region analyzed by EPR spectroscopy for this study. The bottom portion provides the primary sequence of human vimentin within the rod domain presented in the top line. The Jpred4 secondary structure prediction is provided in the second line. The third line provides the heptad assignment of residues by the TWISTER program, with the L1-2 positions designated as (l), and residues part of a hendecad repeat shown in red.

historically defined as the region spanning residues 248–263. We show that the EPR data from this study and previous results (129 positions in total) can be used in conjunction with partial X-ray structural data to construct a molecular model of the complete vimentin rod domain. The resulting model shows a strong correlation with the EPR-derived information on side-chain distances and dynamics. The resulting model facilitates the interpretation of experimental data and provides a platform for molecular dynamic (MD) simulation of IF structural elements and disease-associated mutations. Finally, we demonstrate that the atomistic model in combination with inter-strand spin interactions detected by pulsed EPR, provides a tool for probing higher-order IF assembly. Together these results show that L1-2 is not a random coil joining coiled-coil segments. Instead, our data suggest that L1-2 is an alignment of parallel chains with heterogeneous structure that may play an important role in type III IF filament biology. The data strongly suggest that, throughout the central linker region of the vimentin rod domain, the dimer retains mostly helical, parallel strands. By combining SDSL measurements with MD simulations, we have revised the boundary assignments for the L1-2→rod 2A→L2 regions.

## RESULTS

### EPR Analysis of the L1-2 Region

Figure 1 shows a schematic of the vimentin rod domain, which is dominated by the heptad sequence motifs along with instances of hendecad insertions. However, no clear motif is evident in the region of 249–273 (Figure 1), which includes the region historically designated as L1-2. Much of this region also lacks a helical secondary structure prediction (as determined by JPred4; Drozdetskiy et al., 2015). To determine the structure and dynamics of L1-2, we used EPR spectroscopy of site-directed spin labels of each position beginning in rod 1B through the predicted L1-2 region of vimentin (residues 224–280).

We have previously shown the ability of SDSL-EPR to delineate coiled-coil regions in rod 1 (Aziz et al., 2012; Hess et al., 2004) and rod 2 (Hess et al., 2002, 2006), as well as the lack of coiled-coil structure in the linker 2 (L2) region (Hess et al., 2006). These studies demonstrated that, for samples in the absence of motion (−100°C), a simple empirical parameter for spectral broadening ($d_1/d$; see inset of Figure 2), provides a straightforward method to map side-chain proximity within the vimentin homodimer (Aziz et al., 2009, 2010, 2012; Hess et al., 2002, 2004, 2006), because it provides a model-independent assessment for the proximity of spin-labeled side chains (Likhtenshtein, 1993). For evaluation of a large number of spin-labeled sites, the semi-quantitative relationship of $d_1/d$ and spin label proximity provides a practical method for identifying structural motifs, with close distances (1 nm or less) producing values of $d_1/d > 0.5$, and longer distances (2 nm or more) having values of ~0.33.

The EPR spectra for vimentin positions 224–280 were collected at −100°C (see Figure S1), and the resulting $d_1/d$ values plotted in Figure 2. In the region leading up to the predicted L1-2 domain, there is striking agreement between the amplitude of $d_1/d$ and locations of the *a,d* positions of the heptad repeat (highlighted by blue bars; $d_1/d > 0.44$). It should be noted that the sequence motif underlying the coiled-coil structure of rod 1B is distinct from previous regions examined in that it contains a four-residue extension to the repeat, beginning with position 231. This extended repeat has been designated as a hendecad repeat, with the additional four residues identified as *h-k* (Parry, 2006). However, our EPR analysis (including results described below) suggest that the additional four residues structurally resemble the last four residues of the heptad repeat, and it is therefore more useful to designate them as such. Thus, as shown in Figure 1, residues 238–242 are designated *d-g*, but distinguished as hendecad locations by red font. Using this designation, the data in Figure 2 can clearly identify *a,d* positions for residues 224–250 in both the heptad and hendecad repeats, on the basis of their close proximity within the interface of the vimentin homodimer.

The vimentin fragment PDB: 3UF1 X-ray structure ends with the predicted onset of the L1-2 region (position 251), and a transition from coiled coil to parallel helices at the end of the fragment is apparent in the crystal structure (Aziz et al., 2012). The EPR spectra in Figure 2 indeed show a loss of coiled-coil structure in the region of 251–272, where the sites of close proximity (highlighted in violet) do not correlate to a heptad or hendecad

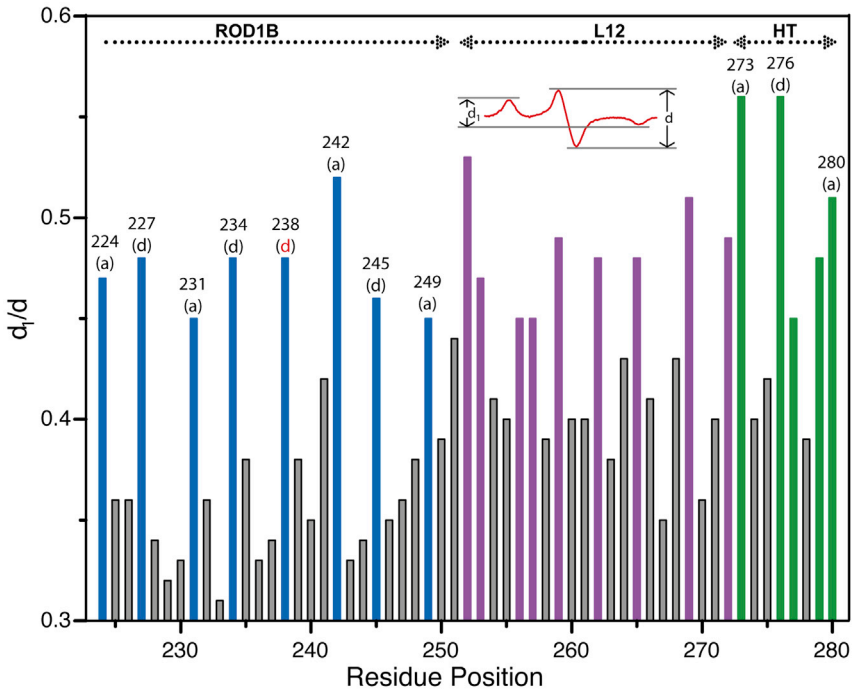



**Figure 2. Plot of the $d_1/d$ Values versus Sequence Position for Residues 224–280 in Human Vimentin**

The heptad residues in rod 1 domain (224–250) with d1/d > 0.44 are highlighted in blue, with the *a,d* heptad position indicated (the *d* position from the hendecad instance is shown in red). Linker 1–2 (L1-2) residues (251–272) with $d_1/d$ > 0.44 are shown in magenta. Sites within the intervening residues ("rod 2A"; 273–280) between L1-2 and L2 with d1/d > 0.44 are shown in green, with the predicted *a* and *d* positions of the heptad indicated.

pattern. These data confirm the transition from coiled-coil structure at the end of rod 1B and are consistent with the L1-2 region arranged as parallel helices (see below).

## Atomic Model of the Entire Vimentin Rod Domain

Because of the difficulty in crystallizing regions containing linkers, as well as the propensity of such regions to adopt non-native structures when occupying the ends of crystallized fragments, we sought to use EPR constraints to build a full-length model of the vimentin rod domain, which can be relaxed in a full-scale manner and evaluated for dynamics. The construction of the full-length rod domain was built using two X-ray structural coordinates: PDB: 3UF1, which encompasses rod 1B, and PDB: 3KLT (Nicolet et al., 2010), which encompasses most of rod 2. The BUILD STRUCTURE module of UCSF CHIMERA was used to build the parallel helices bridging the rod 1 and rod 2 structures. A custom python script enabled the rotation, and the translation of PDB: 3UF1 and PDB: 3KLT to L1-2, while MODELLER was used to structurally optimize the full-length rod domain. In the first step, each piece of PDB: 3UF1, PDB: 3KLT, and L1-2 coordinates were structurally optimized in MODELLER using 22 structurally close positions as determined by the measured $d_1/d$ values from this study (residues 224–280). The positions of the model in this initial optimization maintained a close correlation to the X-ray structural distances between Cα-Cα as well as Cβ-Cβ of the two chains, and served as the physical input distances for the second round of structural optimization in MODELLER. The output from this step resulted in our "unrestrained model." The full-length restrained model underwent an additional optimization using the X-ray structural Cα-Cα and Cβ-Cβ correlation to EPR-derived $d_1/d$ values from throughout the central rod region (a total of 124 out of 129) from this study and previous work (Aziz et al., 2012; Hess et al., 2002, 2004, 2006). Finally, the optimized full-length rod domain model was checked for geometry and scored using PROSA and GA341 scoring functions. An illustration of the model-building process is shown in Figure 3.

A comparison study of the unrestrained model versus the full-restrained model was performed to determine the model quality. For example, without the restraints, the overall root-mean-square deviation (RMSD) between the unrestrained model and the restrained model was 0.51 nm. The unrestrained model was also compared with PDB: 3UF1 (rod 1B), PDB: 3KLT (rod 2B), and PDB: 3TRT (rod 2B) (Chernyatina and Strelkov, 2012) X-ray structures, respectively, with 0.2, 0.67, and 0.34 nm RMSD. While the restrained model to PDB: 3UF1, PDB: 3KLT, and PDB: 3TRT X-ray structures had RMSD values of 0.26, 0.38, and 0.4 nm. Both GA341 and PROSA produced an unrestrained model with a higher deviation from the X-ray structure (*Z* score), while the full-restrained model with a lower *Z* score. We used the restrained model based on closeness to the X-ray structures and with a reasonable overall energy for all the following analyses below. The significance of the EPR restraints on the resulting structure is shown in Figure S2.

In addition to the RMSD analysis, we also performed solvent accessible surface area (SAS) calculations using ALPHASURF (Edelsbrunner and Koehl, 2003) to compare the AB dimer interfaces between X-ray structures. For example, SAS of PDB: 3SSU (144–189; Chernyatina et al., 2012), PDB: 3UF1 (144–249), PDB: 3SWK (153–238; Chernyatina et al., 2012), PDB: 5WHF (153–238; Pang et al., 2018) between the AB dimer resulted in values of 2,140, 5,770, 4,530, and 4,580 $Å^2$, respectively (Table S1). In addition, a value of 5,370 $Å^2$, which is attained for our modeled structure (144–249) correlates well with the dimer interfaces of X-ray structures described above. Concerning the ambiguous structure within the L1-2 and rod 1A/L2 region, we analyzed the dimer interface from the end of rod 1B through the region containing L1-2 and L2 (250–300). The SAS value for this region of our model is 2,370 $Å^2$. Finally, our full modeled structure between 146 and 333 is calculated to have a value of 9,470 $Å^2$ at the interface (Table S1).

## Consistency of the Model with EPR Distance Analysis

The accuracy of the full-length rod domain model can be evaluated from its agreement with X-ray data from vimentin

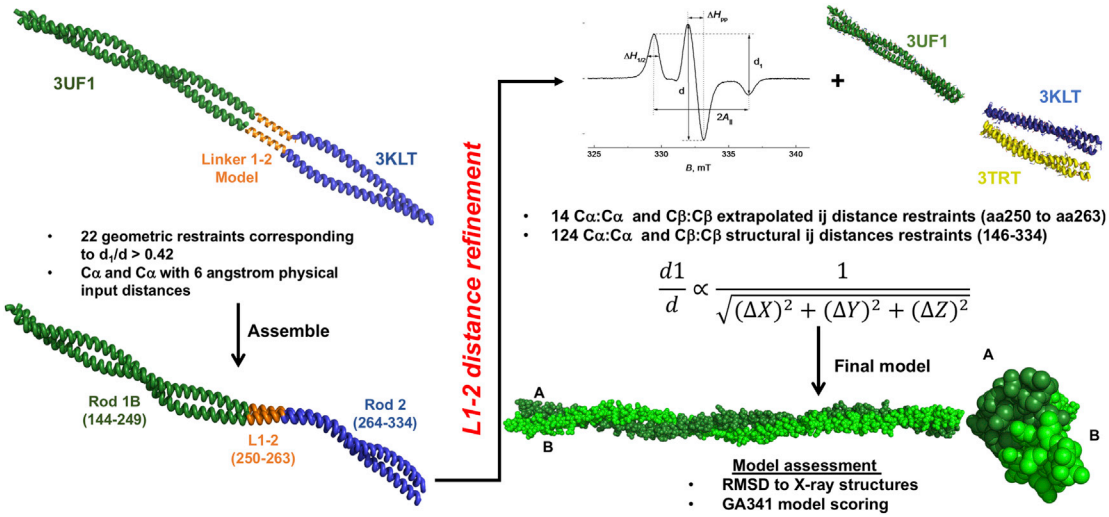



**Figure 3. Model Building of Vimentin Rod 1B (3UF1: 146–249), L1-2 (250–263), and Rod 2B (3KLT: 264–334)**

The construction of the full-length rod domain was assembled using the two protein X-ray structural coordinates: 3UF1 and 3KLT. This includes the vimentin linker 12, using UCSF CHIMERA with rotate and translate of 3UF1 and 3KLT to vimentin linker 1–2, while MODELLER optimized the full-length rod domains. The full-length "restrained model" underwent an additional optimization using the X-ray structural Cα-Cα and Cβ-Cβ correlation to EPR-derived $d_1/d$ values from throughout the central rod region (a total of 124 out of 129) from this study and previous work.

fragments, the EPR data from this study (57 side chains), and EPR results from previous studies. The correlation of experimental $d_1/d$ values spin-spin broadening in the vimentin dimer with the distances between side chains in opposite strands of the dimer is shown in Figure 4. Overall, the model distances between opposing α carbons ($p < 10^{-15}$; Figure 4A), as well as β carbons ($p < 10^{-17}$; Figure 4B), show a strong correlation. For very close Cα-Cα distances (0.8 nm or less), $d_1/d$ is ≥0.44, whereas longer Cα-Cα distances have $d_1/d$ values <0.44. The stronger correlation between the distances between β carbons and $d_1/d$ measurements most likely reflects the ability of the more distal side chain position to reflect the dependence of side distances on orientation. To evaluate how the $d_1/d$ value correlates to the model for different classes of residues within the rod domain, data points in Figure 4 are colored according to their position within the heptad motif (red and blue symbols) or their localization to a linker region (green and violet). Similar plots with the residue position of each point labeled are shown in Figure S3. With respect to the overall correlation, the L1-2 region and the *a,d* positions of the coiled-coil regions are evenly distributed. As a group, the $d_1/d$ values for the *b,c,f,g* positions of the heptad distribute below the fit of the data, although this is expected because of the greater orientation dependence of distances measured between the ends of the spin-labeled side chains. In contrast, the $d_1/d$ values for the previously measured L2 region (Hess et al., 2006) are generally higher than predicted from the model, suggesting that this tightly constrained region may be more relaxed in the model. The relationship between Cα-Cα versus measured $d_1/d$ values, as well as Cβ-Cβ versus $d_1/d$, shows a clear inverse linear correlation and establishes a statistical relationship among physical parameters from two different approaches that can be combined as a general method for structural refinement. In summary, these results support the structural optimization process driven by EPR $d_1/d$ constraints for both X-ray distances and for the modeling of structure in regions that lack a structural template.

## MD of the Constructed Vimentin Model Reflects Unique Structural Features

To further validate the ability of the full-length rod model to recapitulate structural features, we examined the packing and dynamics of two positions within the linker 2 region over a 50-ns time frame. Two positions, 283 and 291, were selected based our previous studies showing these residues possess unique characteristics, including extremely close packing and high thermal stability (Hess et al., 2006). As shown in Figure S4, both 283 and 291 retain a close proximity to the residue on the opposing strand during the simulated dynamics. Likewise, the narrow RMSD of both positions (Figure S4) also reflects the tightly packed environment evident in EPR analysis (Hess et al., 2006). In addition, the molecular model might be useful to explore the basis of this highly stable region. For example, the model reveals an intra-chain salt bridge between K282 and E286 that can be expected to contribute to the high thermal stability of this region. Similarly, the model also reveals that another pair of intra-chain salt bridges between E289 and K292 is expected to contribute to the high thermal stability as well as restricting the side-chain motion of Y291. These restricted dynamics in the model clearly identify the uniqueness of the L2 region that we have previously determined experimentally (Hess et al., 2006).

## Correlation of MD to Spin Label Motion

To evaluate the dynamics of residues 224–280 in vimentin, we collected EPR spectra at room temperature of individually spin-labeled vimentin, as assembled into protofilaments (Mucke et al., 2004). The line widths of the solution spectrum of a

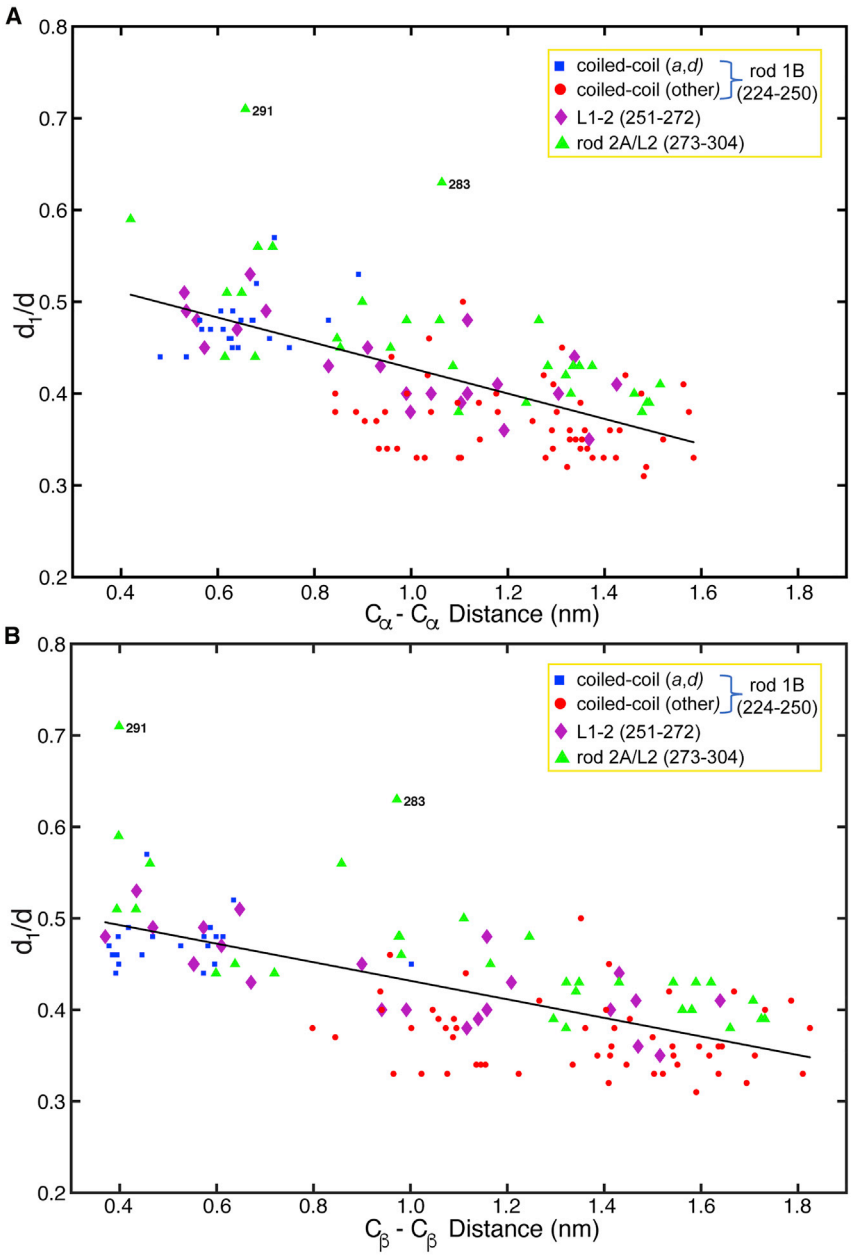



**Figure 4. Correlation between Distances in the Model and EPR Results**

For each sequence position in the vimentin dimer, the distance between the A and B chains in the model are plotted according to its EPR-measured $d_1/d$ value. The correlation of $d_1/d$ to both the Cα-Cα distance (A) and the Cβ-Cβ distance (B) is shown. In each graph, the $d_1/d$ values for 129 spin-labeled positions are plotted against side-chain distances with symbols indicating structural domain assignment. For coiled-coil regions, distances are defined by the X-ray structures 3UF1 and 3KLT. The distances for the residues in linker regions are obtained from the optimized model. The Pearson correlation coefficient indicates a strong linear relationship between Cα-Cα and $d_1/d$ ($p < 10^{-15}$), and the correlation is even stronger when using the Cβ-Cβ distance ($p < 10^{-17}$), which can better account for side-chain projections. Positions 283 and 291 are highlighted due to their unique stability properties (Hess et al., 2006). p values were determined using the fitlm linear regression model in MATLAB.

spin-labeled side chain is dependent on both backbone and side-chain dynamics (Altenbach et al., 2015; Klare, 2013), and therefore provides a direct method for identifying structural motifs in vimentin (Aziz et al., 2012; Hess et al., 2002, 2004, 2006, 2013). Figure 5A shows the room temperature EPR spectra collected for the region representing the final heptad repeat of rod 1 (residues 242–250), with the spectral amplitudes normalized to the same number of spins. Qualitative examination of the spectra reveals three positions (242, 245, and 249) that are extremely broadened. Consistent with previous studies (Aziz et al., 2012; Hess et al., 2002, 2004, 2006, 2013), the broadening at these sites can be attributed to their location at the interface of the coiled coil (i.e., *a* or *d* positions within the heptad repeat), which imparts both motional restriction and dipolar coupling between side chains in close proximity. A similar correlation between spectral line shape and heptad position (including the hendecad *d* assignment of position 239, see above) is evident for the upstream region of rod 1B (residues 224–241; see Figure S5).

In contrast to the predicted rod 1B region (residues 224–250), the side-chain dynamics of residues in the L1-2 region (251–272) do not display a pattern of coiled-coil structure (Figure S5). Rather, multiple positions display spectral broadening, indicative of a rigid and close arrangement of the dimer strands in this region. Notably, the extent of broadening in the room temperature spectra closely follows the broadening seen in the spectra of frozen samples (expressed as $d_1/d$, see Figure 2). Remarkably, two positions (264 and 265) display two clearly resolved dynamic populations, a feature not present in any of the other 127 residues examined within the vimentin rod region. Whether this unique feature arises from a unique side-chain rotamer distribution in each strand, or distinct dynamics within each strand is discussed below.

While multicomponent fitting of solution EPR spectra provides a quantitative assessment of spin label correlation times, the relative dynamics of spin-labeled sites can be determined by ($\Delta H^{-1}$), a simple model-independent index for the degree of side-chain order (Altenbach et al., 2015). The map of side-chain dynamics determined by $\Delta H^{-1}$ was then compared with the computed B factors of side chains obtained from a 50-ns MD simulation of the constructed model. The results are shown in Figure 5B. Similar to the results based on distance between opposite side chains (Figure 4), there is a strong correlation between the EPR-measured dynamics and

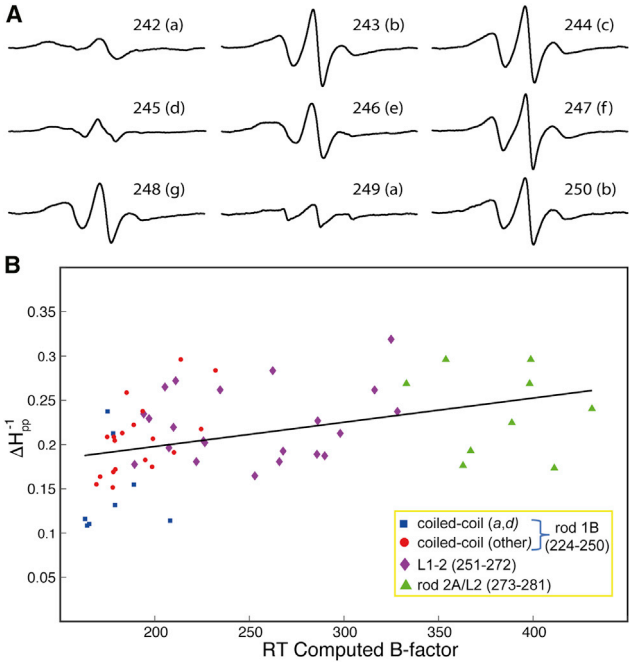



**Figure 5. Relationship of Computed Dynamics to the Dynamics of Spin-Labeled Side Chains**

(A) Solution EPR spectra of vimentin acquired at room temperature. Scan width is 100 G. The letters in parentheses represent the location of the residue within the heptad.

(B) The relationship between the $\Delta H_{pp}^{-1}$ (see Figure 1) values of the solution EPR data to the room temperature (RT) computed B factor parameters of the model. Fifty-eight data points of $\Delta H_{pp}^{-1}$ were compared against the computed B factor of the restrained model over a 50-ns MD simulation to determine a relationship between the two physical parameters. The symbols for each position are distinguished according to structural domain assignment.

the side-chain order calculated through MD simulation of the model, a finding that further supports the utility of the model in understanding vimentin structure and dynamics. For example, the analysis reveals that the L2 region has greater thermal motion compared with L1-2. Moreover, larger concerted motions can be visualized from the MD simulation of the constructed vimentin model, such as the visualization of the greater flexibility in the rod 2 domain compared with the rod 1 domain (see Video S1).

The comparison of simulated dynamics of native residues to measured side-chain dynamics is facilitated by the preferred rotamer of the 1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl methanethiosulfonate (MTSL) spin-labeled side chain on helical structure (Polyhach et al., 2011). To verify this predicted behavior, we performed a single MD simulation of the spin-labeled side chain at position 224 to evaluate its probable conformer distribution on a vimentin helix. As shown in Figure S6, the spin-labeled side chain populates a densely compacted space, indicative of a preferred rotamer. Moreover, the spatial distribution of the N-O moiety is largely determined by the $\chi_1$-$\chi_5$ angle of the MTSL-labeled Cys side chain. This rotamer of the MTSL-labeled Cys has been identified as a favored conformation along helical structure (Polyhach et al., 2011).

## Analysis of L1-2 Dimensions by Pulsed EPR

To differentiate between the two hypotheses that L1-2 is largely helical as our EPR and model predict, or dynamic as hypothesized by Parry and Smith (2010), we prepared vimentin molecules with a deuterated isoleucine combined with a targeted spin label to detect close interactions within the 256–260 region by electron spin echo envelope modulation (ESEEM) spectroscopy. This technique uses both a specifically deuterated amino acid and a spin-labeled position to report unambiguously whether the two labeled sites are within 0.8 nm of each other (Liu et al., 2016). For $\alpha$ helices, this condition is met for three or four residues removed from the deuterated position. In contrast, for $\beta$ strands, this condition is met for residues two positions removed from the deuterated residue. To apply this approach to the middle of the L1-2 region, *E. coli* transformed with single-Cys mutations in vimentin were grown in media containing deuterated isoleucine and thus produce proteins isotopically labeled at I256. For each of the three single-Cys mutations (at positions 258, 259, or 260) containing the deuterated Ile, the expressed vimentin was purified, and spin labeled with MTSL. To minimize interstrand couplings, the labeled protein was diluted with unlabeled vimentin in urea (at a ratio of 1 labeled:5 unlabeled). The spin-diluted samples were then dialyzed against low-ionic strength tris buffer and analyzed by ESEEM. Surprisingly, ESEEM couplings are observed for all three spin-labeled positions (Figure 6A). This suggests either a mixture of $\beta$ sheet and $\alpha$-helical structure, or a distinct backbone fold where position 256 would maintain a close proximity to each of the three spin-labeled locations (258, 259, and 260).

Although the MD simulations of the constructed model do not show evidence for $\beta$ strand formation, the helical distance within this region of L1-2 is distorted relative to a typical $\alpha$ helix. In fact, over a 50-ns simulation, the distance of the C$\beta$ position on 256 to the C$\beta$ positions in each of the three spin-labeled sites falls within the 0.8 nm threshold for ESEEM (Figure S7). Thus, our model suggests the successive ESEEM signals arise from an unusual helical fold rather than the presence of $\beta$ secondary structure.

To test the dimensions of the L1-2 region over a larger scale, we used a second pulsed EPR approach capable of detecting spin couplings out to 5 nm or more. This method, double-electron electron resonance (DEER) (Jeschke, 2012), we made double-Cys mutants within vimentin and labeled both positions with the MTSL spin label. The double-labeled protein was then spin-diluted (to minimize intermolecular spin couplings) with unlabeled wild-type (WT) vimentin at a ratio of 1:5 (labeled:unlabeled), assembled and measured for inter-spin distances by DEER. We generated three spin-pair combinations for DEER: (1) positions 225–240 to serve as a reference for a known helical segment in rod 1B; (2) positions 244–265, encompassing the end of rod 1B and the first 15 residues of L1-2; (3) positions 250–265, encompassing the first 15 residues of L1-2.

The results of the DEER measurements are shown in Figure 6B. Due to the spin dilution and the difficulty of concentrating vimentin without inducing aggregation, the signal-to-noise ratio of the data were not ideal. However, as demonstrated by the control rod 1B sample (225–240), the DEER result matches the distance calculated from the 3UFI X-ray structure (Table S2). In contrast, the DEER analysis of the 244–265 pair does not produce a single

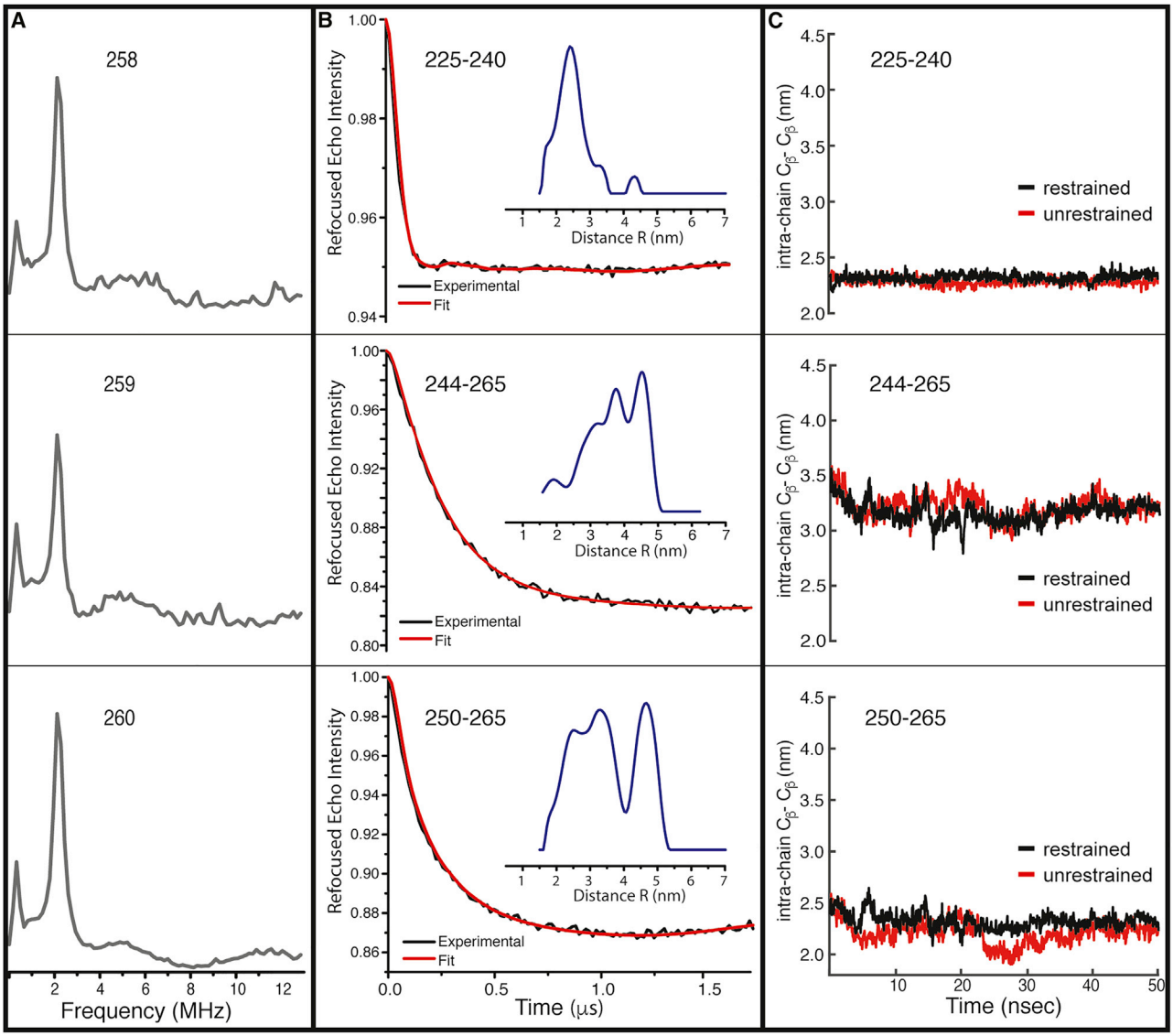



**Figure 6. Pulsed EPR Data of Vimentin L1-2 Region**

(A) The three-pulse ESEEM frequency domain data of vimentin containing a deuterated ($d_{10}$) Ile256 paired with nitroxide spin labels placed at positions 258, 259, or 260. Weak dipolar couplings between the $^2$H nuclear spin and the nitroxide are detected for all three spin label locations (i+2, i+3, and i+4). Time domain data and Cβ-Cβ distance calculations from MD simulation are shown in Figure S7.

(B) The four-pulse Q-band DEER data for the indicated spin pairs introduced into vimentin. Decays represent the background-subtracted dipolar evolutions of the spins, with the distance probability distributions from Tikhonov regularization given in the inset. To limit intermolecular interaction of spin labels, each double-labeled protein was spin-diluted with unlabeled WT vimentin at a ratio of 1:5 (labeled:unlabeled).

(C) The average distance between Cβ positions in model over 50 ns of MD simulation.

major peak. In this case, three major distance peaks are obtained, 3.2, 3.7, and 4.5 nm (Table S2). The 3.2-nm DEER distance agrees well with the Cβ-Cβ distance measured in our model (3.2 nm). The source of the additional DEER peaks observed for the 244–265 pair is discussed below. The DEER results for the 250–265 pair are also complicated by multiple distances, with three major DEER distance distributions (Figure 6B). Of the three peaks, the distribution centered at 2.5 nm is consistent with the intramolecular distance between the positions in our model (2.3 nm; Table S2), and therefore consistent with a distorted helical structure for the L1-2 region.

The additional DEER peaks for the 250–265 and 244–265 pairs could arise from a mixture of structural conformations in the L1-2 region. However, the presence of three distance peaks in the 250–265 pair is better explained by an A12 alignment of assembled vimentin units centered near position 250. In the A12 alignment, adjacent dimers are aligned in an antiparallel manner with overlap in the L1-2 region (Steinert et al., 1993). Thus, it is not surprising that positions in the vicinity of L1-2 experience intermolecular (inter-strand) interactions with spin labels located on adjacent dimers. In fact, when combined with the full-length molecular model of vimentin, a collection of intermolecular DEER

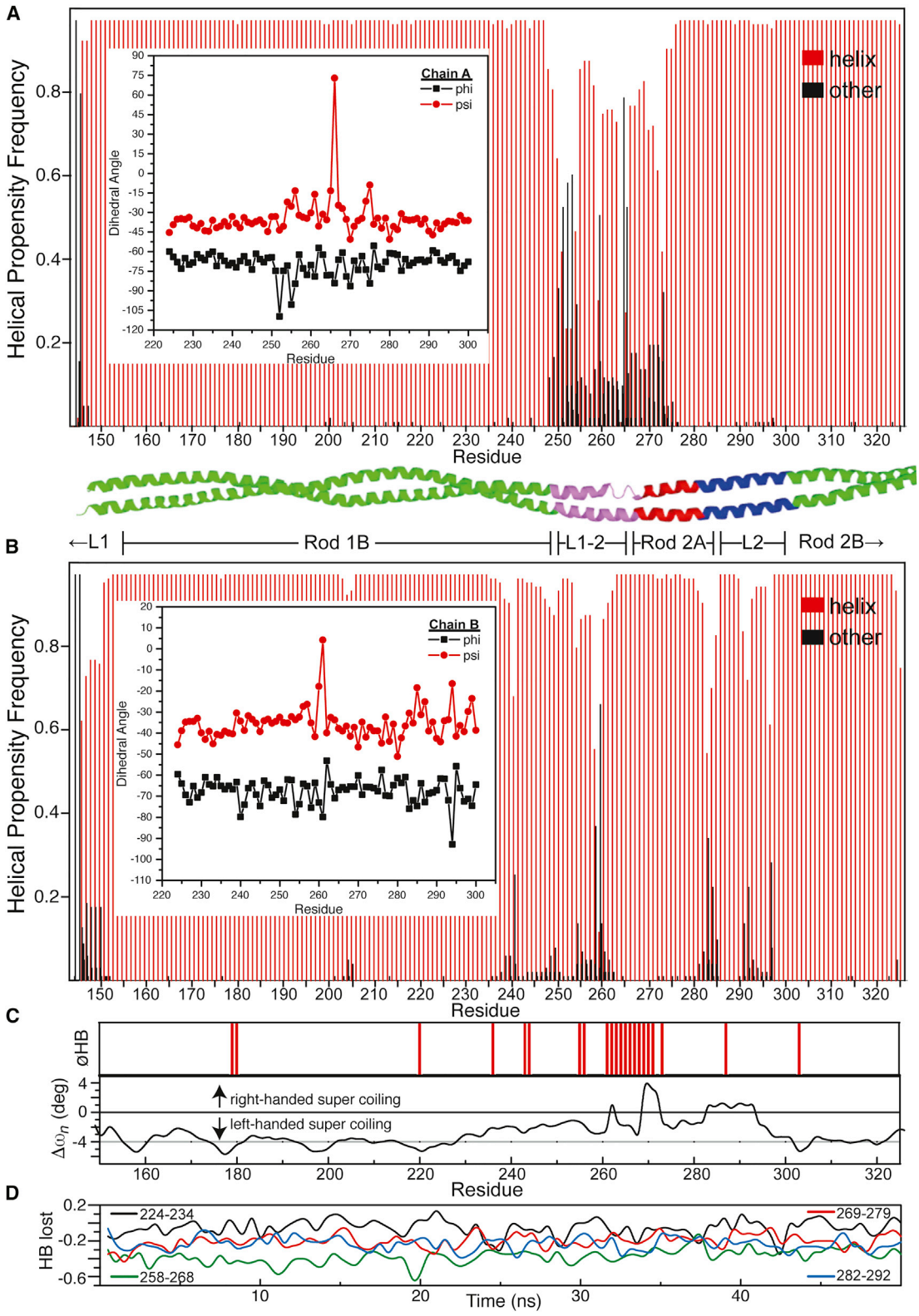



**Figure 7. A Quantitative Definition of Structural Domains within Vimentin**

(A and B) The mixed α-helical distortion was structurally characterized based on 100 frames of 50-ns MD trajectory, and coordinates entered into the DSSP program. The α-helical frequency was computed for every position in the vimentin sequence range of 144–325 and plotted for chains A and chain B

*(legend continued on next page)*

distances provides a method for modeling the assembly of vimentin and other IFs. In this regard, we used the multiple DEER distances for the 250–265 and 244–265 pairs to model an A12 alignment of vimentin dimers (chains *AB* aligned with chains *CD*). The resulting model is shown in Figure S8, with an assignment of the model Cβ-Cβ distances to a measured spin-pair DEER distance given in Table S3. The resulting intermolecular Cβ-Cβ distances in the modeled A12 alignment are close to the distances reported by DEER. The exception is the Cβ-Cβ distance of 2.0 nm between position 250 in chains *B* and *C* in the model. The nearest DEER peak for the 250–265 sample is 2.5 nm. However, the lack of a DEER peak in the 2-nm range is not surprising. Distances on the order of 2 nm or less are difficult to detect with DEER, because the couplings become too intense to be observed within the pulse bandwidth. Alternatively, the 250B-250C species could be contributing to the 2.5 nm DEER peak, implying the Cβ-Cβ distance in the model underestimates the distance between the ends of the spin labels as measured by DEER.

Although the modeled A12 alignment adequately accounts for the additional DEER distributions, the amplitudes of the additional peaks are larger than expected for samples that have been spin-diluted to minimize intermolecular interactions. For example, within a dimer, the probability of intermolecular interaction is 20% given the 1:5 (labeled:unlabeled) spin dilution. However, the probability for intermolecular interaction with labels located in an adjacent dimer in the A12 alignment is higher, since spin-coupling can be achieved with spin labels located in either chain. In addition, for labels located at position 250, the possibility of coupling to labels located on more than one tetramer exists. The idea is illustrated in Figure S9, which shows a schematic of possible interactions of vimentin tetramers in the A12 alignment (based on the elongated IF cross-section; Sokolova et al., 2006). Finally, with respect to the relationship of distribution peak amplitudes to the population of spin pairs, the DEER peak assigned to the intramolecular coupling is broader for both the 244–265 and 250–265 spin pairs. As detailed below, the chains of the vimentin dimer show structural heterogeneity in the L1-2 region despite having the identical sequence. Thus, the 244–265 and 250–265 spin pairs may report different distances, depending on the conformational state of the chain. Indeed, dimensions from the model suggest a significant length difference between the two chains. Measuring the region of residues 250–300, the end to end distances are 7.6 and 7.9 nm for chains A and B, respectively.

## MD Simulation Identifies Structural Heterogeneity among the Paired L1-2 Strands

Vimentin L1-2 is historically defined as positions 248–263. However, MD simulations define this region as a helical distortion that appears to propagate from positions 251 to 272. This mixed helical distortion was further characterized structurally based on 100 frames of a 50-ns MD trajectory using the DSSP program. The structural coordinates derived from the MD trajectory were entered into a DSSP program, and the α-helical frequency was computed for every position in the vimentin sequence range of 144–334 in both chain A and chain B. Interestingly, the α-helical distortion frequency extends out from positions 251 to 272 with the structural heterogeneity predominately for chain A (Figure 7A) compared with chain B (Figure 7B). For the region encompassing L2, chain B shows smaller helical distortions in the sequence range of 282–300. Furthermore, although the average ϕ-ψ values (−65° and −40°, respectively) for the L1-2 region resemble a typical α helix, the uniqueness of this region is also revealed by alterations in ϕ and ψ angle distribution (Figure 7B, inset). According to these parameters, residues D264 and T266 for chain A show the largest dihedral deviation of helical character and appear to have a propensity for β strand character. Compared to a coiled-coil region, the total energy of the L1-2 region is slightly higher when calculated by a molecular mechanics force field (Maier et al., 2015), with an average value of −182 kcal/mol for chains A and B in the relaxed structure compared with −191 kcal/mol for the coil-coiled region of residues 224–234.

L1-2 is readily identified by its lack of a coherent heptad pattern. Such transitions in coils can manifest as phase shifts, imparting strains that affect the degree and handedness of supercoiling (Brown et al., 1996; Lupas and Gruber, 2005). To evaluate the degree and direction of coiling, the phase yield per residue ($\Delta\omega_n$) (Strelkov and Burkhard, 2002) was calculated along the optimized vimentin structure. A canonical left-handed coiled coil has a $\Delta\omega_n = -4°$, whereas a helix without twist has a value of zero, and right-handed supercoiling has positive values of $\Delta\omega_n$. As shown in Figure 7C, the unwinding of rod 1B begins with the hendecad insert at position 238 and continues through L1-2 and into rod 2A, where a right-handed twisting force is predicted. The strain at the end of L1-2 (where we also see structural heterogeneity between the chains) is also reflected in the number of backbone H-bonds lost in the optimized structure (Figure 7C, top panel). To obtain a comparative analysis of strain over time, the frequency of H-bonds lost was calculated for the distinct regions over 50 ns of MD simulation (Figure 7D). Again, L1-2 shows the highest frequency of H-bond disruption, whereas the coiled-coil segment of rod 1B experiences the lowest frequency.

These results provide a quantitative basis for defining the structural ambiguity of the region spanning L1-2, rod 2A, and L2. Taken together with SDSL measurements, the conformationally unique region of L1-2 is best defined as residues 251–268, positions 269–281 encompassing rod 2A and L2 as residues 282–300. In addition to better delineating the boundaries of these sub-regions within the central region of vimentin, these analyses along with SDSL provide insight onto the relative order within each subdomain. In particular, the sub-region of L1-2 distinguished for its chain heterogeneity and L2 for its thermal stability.

(A and B, respectively). Chain A shows greater helical distortion in the L1-2 region, whereas chain B shows greater helical distortion in the rod 2A/L2 region. The uniqueness of L1-2 is further identified by dihedral angle variation (insets of A and B).

(C and D) The coiled-coil phase yield ($\Delta\omega_n$) per residue as a function of residue number for the restrained vimentin model calculated by TWISTER (C). The phase yield value of an ideal left-handed coiled coil (−4) is indicated by the gray line. The location of residues lacking backbone H-bonds (øHB) is shown in the upper sub-panel. A plot comparing the frequency of H-bonds lost for the distinct regions is shown in (D). Plotted are the frequencies for the rod 1B (224–234), L1-2 (258–268), rod 2A (269–279), and L2 (282–292) regions over 50 ns of MD simulation.

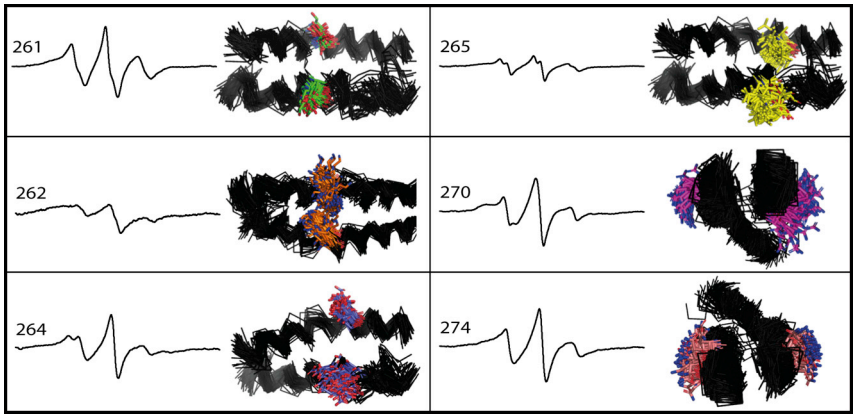



**Figure 8. Side-Chain Motions Show Distinct Two Conformational Motions that Correlate with Two-Component EPR Spectra**

Side-chain dynamics were structurally characterized based on 100 frames of 50-ns WT vimentin MD trajectory.

As discussed above, we observed a distinct structural diversity of L1-2 despite the sequence of homodimer being identical. To probe this further, we ran 50 ns of MD simulation for L1-2 dynamics. Throughout this time the heterogeneity of the two chains is maintained; however, a switching between alternate states is observed (one chain is helical, the other distorted), showing evidence of concerted dynamics (Video S2). In addition, the EPR data demonstrate that some sites in L1-2 supports the notion of discrete states for chains A and B (Figure 8).

### Using the Model for Comparative Analysis of L1-2 Sequences and Mutations

We also extended the structural heterogeneity of L1-2 for two other type III IFs (GFAP and DESMIN) by performing 50 ns of L1-2 region MD simulations. Specifically, MD simulation of L1-2 A and B chains in WT vimentin reveals a switching of helical distortion between the chains, revealing a dynamic asymmetry (Figure 9B). This observation was the same for both GFAP and DESMIN. In all three cases (GFAP, DESMIN, and VIMENTIN L1-2 region), one backbone chain is well ordered, while the other is distorted.

Using the homodimer model of vimentin, we generated a homology model of the WT keratin14/5 heterodimer as well as a mutated form of the epidermolysis bullosa simplex (EBS), and Alexander disease into the vimentin homodimer. D257Q was substituted *in silico* for the vimentin model and simulated for 50 ns to mimic the Alexander disease mutation. We observed a distorted L1-2 in which bent helices were unable to be rescued throughout the 50 ns. Furthermore, mimics of disease-associated mutation induce global changes in the IF structure, suggesting that the L1-2 region resides near a conformational threshold (Figure 9C). Similarly, both D259G and K262C mutations, which mimic the EBS disease mutant form of the vimentin model, were performed for 10 ns of MD simulation. Much like the D257Q mutant vimentin, we observed distorted L1-2 bent helices in the D259G and K262C mutations (Figures 9D–9F).

Mean interaction energy (kcal/mol) between chain A and chain B was computed based on the MD simulation runs (Table S4). We observed that WT vimentin and keratin have similar mean interaction energies of −84.9 and −88.7 kcal/mol, respectively. Interestingly, D259G has a mean interaction energy of −416.1 kcal/mol; D257Q has mean interaction energy of −332.3 kcal/mol; and K262C has mean interaction energy of −73.5 kcal/mol.

## DISCUSSION

Our analysis of L1-2 and modeling of the completed assembly reveals that the middle of vimentin differs from the historically predicted structure and provides revised sequence delineations for the structural designations within this region. Our data confirm X-ray crystallography showing that the carboxylic end of rod 1B as well as rod 2A through L2 adopt a parallel helices structure instead of a coiled-coil structure (Aziz et al., 2012; Chernyatina and Strelkov, 2012; Nicolet et al., 2010). For the region encompassing rod 2A and L2, the parallel helices feature has been identified as a "paired bundle" (Chernyatina et al., 2015). With the spectroscopic analysis of L1-2 completed, the central rod domain of vimentin can be described in a simpler manner. Instead of a structure composed of four coiled-coil domains and three linker domains (rod 1A, L1, rod 1B, L1-2, rod 2A, L2, and rod 2B), our data are consistent with a tripartite central rod domain: a central parallel helices domain flanked by left-handed coiled-coil domains. It should be noted that, within this central linker region, there are distinct levels of order roughly corresponding to the L1-2, rod 2A, and L2 designations (especially the chain heterogeneity of the L1-2 portion, and the highly stable nature of the L2 portion). Based on these findings, we present a revised schematic map of the vimentin structural domains in Figure 10, which includes the locations of structural motifs common to IF proteins.

Our designation of rod domains is based upon direct evidence for left-handed coiled-coil structure established by crystal structures or detection of the heptad repeat by SDSL-EPR. However, insertions and deletions into the heptad can either loosen or tighten supercoiling within coils (Lupas et al., 2017), such as the hendecad repeat near the end rod 1B (Figure 1) that serves to unwind the helices. SDSL-EPR data detect the hendecad repeats at the end of rod 1B and the region encompassing L2. No coil pattern is evident for the L1-2 region, where SDSL-EPR, pulsed EPR, and simulations are consistent with a metastable structure. In addition, SOCKET (Walshaw and Woolfson, 2001) analysis for "knobs-into-holes" over the assembled structure assigns coiled-coil geometry for residues 154–245 and 305–330, but not in the region spanning L1-2 through L2. Although *a/d* proximity for positions 273 and 276 in rod 2A is detected by SDSL-EPR, this segment may be too short to impart the twist commensurate with knobs-into-holes packing.

Although X-ray crystallography has been extremely useful for obtaining structural information on IF proteins, fragmented and modified vimentin peptides can adopt non-native or disordered structure within the unit cell. For example, PDB: 3KLT (Nicolet et al., 2010) displays high thermal motions at its

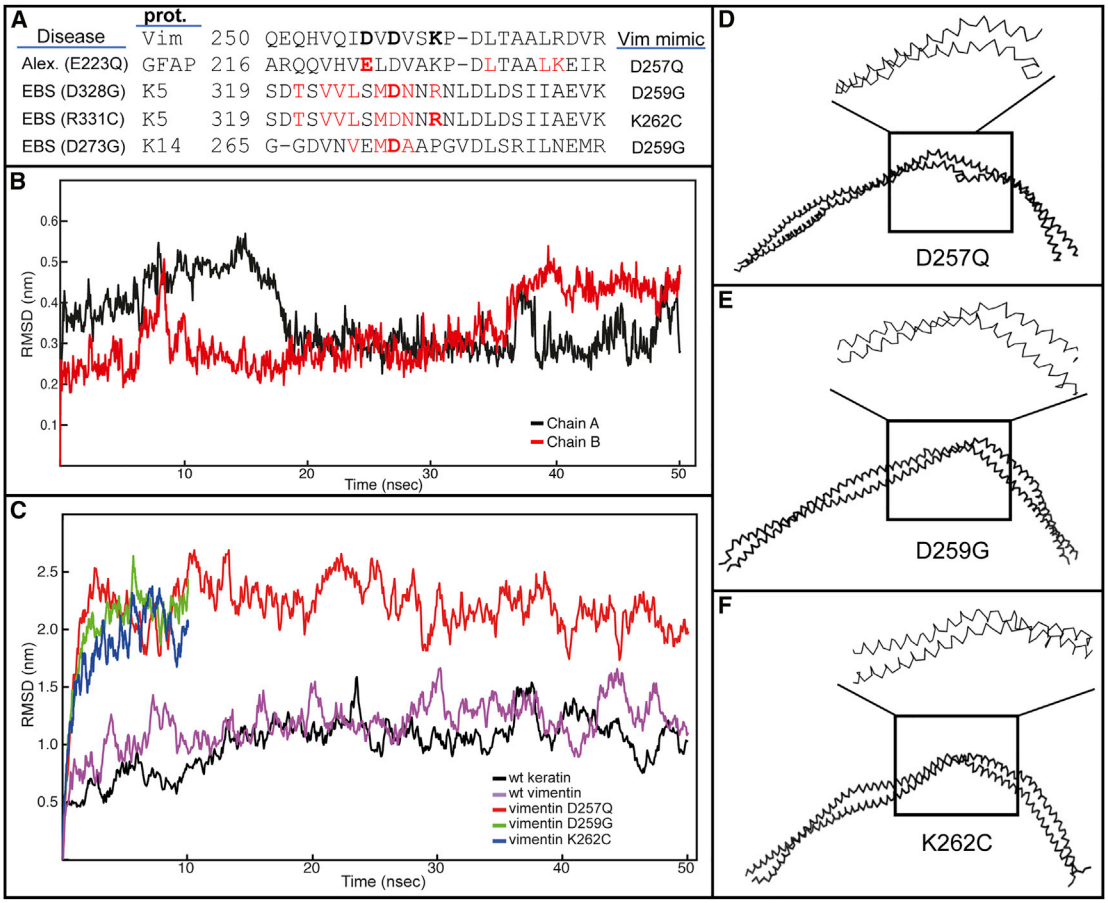



**Figure 9. Consequences of Mutations in the L1-2 Region of IF Proteins Explored by MD Simulation**

(A) Sequence alignment of L1-2 region of human IF proteins. Mutations within L1-2 in GFAP and K5/14 that have been linked to human disease are shown in red. Positions used for disease mimic in vimentin are shown in bold.

(B) MD simulation of L1-2 A and B chains in WT vimentin reveals a switching of helical distortion between the chains, revealing a dynamic asymmetry.

(C) Mimics of disease-associated mutations induce global changes in IF structure, suggesting that the L1-2 region resides near a conformational threshold. RMSD calculations were performed for the D257Q, D259G, and K262C mutations and compared with WT keratin 14/5 and vimentin.

(D–F) Snapshots of the MD simulations show that the L1-2 region becomes distorted and bent upon introduction of the disease-causing mutations.

N-terminal end (residues 263 through 268), resulting in a mean crystallographic B factor of 58. The same structure also contains uncertainty at position 334. Another example of structural uncertainty is observed in PDB: 3SSU (Chernyatina et al., 2012), which lacks reliable information for positions 144 to 148 (a mean crystallographic B factor value of 131). In the PDB: 3SWK (Chernyatina et al., 2012) structure for the 153–238 vimentin region, the N-terminal position has a mean B factor of 66. Finally, in the overlapping region beyond position 221 in the PDB: 5WHF, PDB: 3UF1, and PDB: 3SWK, vimentin crystal fragments deviate significantly (RMSD > 3.0). Within this region, the coordinates of our final model reside near the average positions of these three structures. SDSL-EPR is therefore very complementary to crystallographic efforts, in that it can clarify local ambiguities in the structure.

X-ray crystallography of a vimentin fragment including the rod 2A-L2 region revealed two pairs of widely spaced, parallel helices interlaced into tetrameric structure with apparently normal coiled-coil domains forming downstream (Nicolet et al., 2010). Based on this result, rod 2A through L2 was suggested to adopt a continuous region of straight helical chains running in parallel, consistent with previous EPR data of L2 (Hess et al., 2006). In the case of the L1-2 region, SDSL-EPR shows that the PDB: 3TRT (Chernyatina et al., 2012) structure is influenced by the engineered disulfide at the fragment's N-terminal end (residues 261–269). After position 270, the disulfide shows no further perturbation on the vimentin structure, and there is good agreement between SDSL and the PDB: 3TRT structure. Finally, there is a significant divergence from structure PDB: 3UF1 (Aziz et al., 2012) to structure PDB: 5WHF (Pang et al., 2018) in the sequence range of 221–238, suggesting that dynamics within this region produces multiple states. We propose an alternative approach in comparing these two structures by superimposing it on our dynamic model, which possesses both dynamic spin label restraints at particular positions as well as using the X-ray structural distances with energy minimization.

Although the initial building of our model was based upon the PDB: 3UF1 (Aziz et al., 2012) and PDB: 3KLT (Nicolet et al., 2010) structures, our final model complements information reported by

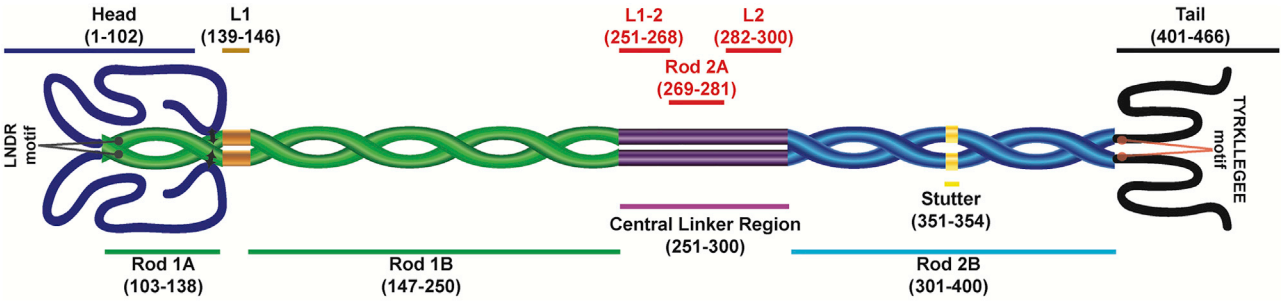



**Figure 10. Revised Structural Map of the Central Linker Region in Vimentin**

The revised boundaries of linker 1–2 (L1-2), rod 2A, and linker 2 (L2) are indicated. These compare with the historical designations (cf. Herrmann and Aebi, 2004) of 248–263 for L1-2, 264–282 for rod 2A, and 283–290 for L2. EPR restraints provided in this study span from the middle of rod 1B through rod 2A. Previous EPR studies revealed that L2 is comprised of ridged straight helices (Hess et al., 2006). The coiled-coil structure of the rod 1A/B and rod 2B segments is indicated, with definitions established by previous EPR and X-ray crystallography studies: rod 1A (Aziz et al., 2009; Meier et al., 2009), rod 1B (Aziz et al., 2012; Chernyatina et al., 2012), and rod 2B (Hess et al., 2002, 2006; Nicolet et al., 2010; Strelkov et al., 2002). Also noted is the LNDR motif at the beginning of rod 1A, a site in IF proteins associated with congenital skin disease (Hess et al., 2005; Omary, 2009), as well as the stutter sequence in rod 2B (Brown et al., 1996; Strelkov et al., 2002). The structures of the head and tail domains remain unknown; however, aspects such as the localization of the head to rod 1A (Aziz et al., 2009) (black double arrows) and the level of structural order in the tail (Hess et al., 2013) have been determined. The YTRKLLEGEE motif at the beginning of the tail is also noted, as it is conserved among IF proteins (Herrmann et al., 2000; Parry and Steinert, 1999) and shows a close proximity among the acidic residues in each strand (Hess et al., 2013).

other crystallography studies on vimentin fragments. As noted above, there is ambiguity in aspects of these structures, such as the significant divergence in the region of residues 221–235 between the recently published PDB: 5WHF (Pang et al., 2018) and PDB: 3UF1 X-ray structures. Our results demonstrate the utility of EPR-derived constraints in model optimization to clarify structures and provide dynamic information that is absent in crystal structures. When excluding ambiguous regions of X-ray structure (due to missing residues in the X-ray structures or having high B factor values, such as PDB: 3KLT and PDB: 3SSU) from our model, we find a strong agreement with other vimentin crystal structures. For example, superimposition of PDB: 3UF1 (residues 146–220) and PDB: 3KLT (residues 269–333) via CEALIGN on chains A and B of our final model results in RMSD values of 0.24 and 0.35 nm, respectively. In addition, our modeled structure for the vimentin rod 1A region resembles the PDB: 5WHF structure of this region reported by Pang et al. (2018), where a superimposition of positions of 154–220 in chains A and B via CEALIGN on our final model results in an RMSD of 0.3 nm. Other structures, such as PDB: 3SSU (Chernyatina et al., 2012) (residues 149–187), PDB: 3SWK (Chernyatina et al., 2012) (residues 154–237), PDB: 3S4R (Chernyatina et al., 2012) (residues 149–189), and PDB: 3TRT (Chernyatina et al., 2012) (residues 270–333) superimpose on our final model results in RMSD values of 0.13, 0.3, 0.18, and 0.37 nm, respectively.

EPR and MD data of L1-2 suggest distinct conformations among the chains in the dimer (helical and distorted helix). Intriguingly, the two protein strands have the ability to "flip" in a concerted mechanism from one conformation to another. We view this dynamic as a plausible explanation for the hypothesis of Parry and Smith (2010), suggesting that this region first adopts a $\beta$ structure, and that a $\beta$ to $\alpha$ switch occurs during assembly of the L1-2 helix. Modeling also suggests that paired H-bond networks around the vimentin L1-2 may stabilize the mixed helical nature of the region. In particular, 255 OE1 to 255 NE2, as well as 255 NE2 to 259 OD1, remain extremely stable through the MD simulation trajectory. This alternate conformation of identical protein sequences in vimentin was also seen in the structure of rod 1B assembled into an A11-like tetramer (Aziz et al., 2012). The chain distinction within homopolymeric vimentin may be an inherent property of IFs, given that this feature is expected among heteropolymeric IFs. Furthermore, the strain within the L1-2 structure suggests that this region of vimentin may be structurally sensitive to mutation. As such, a D257Q substitution to mimic a mutation in GFAP, which leads to Alexander disease, was probed by MD simulation. The resulting bent structure of L1-2 remains stable throughout the 50-ns simulation. It is therefore reasonable to postulate that the stability of the native L1-2 region resides near a conformational threshold, with the D257Q point mutation switching the region into a relaxed, bent structure that would impede filament assembly.

The determination of the structure of the central rod domain of vimentin using spectroscopic techniques, followed by the creation of a valid molecular model, opens the door for several future studies. We previously published spectroscopic structural data for the vimentin head and tail domains, regions for which no X-ray crystal structures have been solved. Using the modeling techniques described here, it is possible that molecular structures for these domains could be created. With or without models of the head and/or tail domains, the existence of a molecular model of the central rod domain opens up the study of higher-order vimentin assembly using *in silico* modeling of A11, A22, and A12 interactions, because vimentin molecules pack in different ways to generate a 10-nm filament. As exemplified in this study, spin labels in the L1-2 vicinity support the existence of the A12 tetramer structure.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY

- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Site-Directed Mutagenesis and Purification for Spin Labeling
  - In Vitro Filament Assembly and Electron Microscopy
  - EPR Spectroscopy of Site-Directed Spin Labels
  - Model Building of L1-2 Linker 250-262 (Chain A) and 250-264 (Chain B)
  - Structural Alignment of 3UFI and 3KLT PDB Coordinates to the L1-2 Linker
  - Cα–Cα Spatial Restraints of Chain A to Chain B for Parallel Helix Refinement via MODELLER
  - Molecular Dynamics Studies
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND CODE AVAILABILITY


## SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.str.2019.07.011.

## ACKNOWLEDGMENTS

The authors acknowledge grants EY015560, NEI Core Facilities grant P30-EY012576, and NIH Research Facilities Improvement Program Grant C06 RR-12088-01 to P.G.F. G.A.L. would like to acknowledge a NIGMS/NIH R35 GM126935 award. The pulsed EPR spectrometer was purchased through funding provided by the NSF (MRI-1725502).

## AUTHOR CONTRIBUTIONS

Conceptualization, J.C.V., J.F.H., D.D.G., and P.G.F.; Methodology, J.C.V., J.F.H., and D.D.G.; Software, D.D.G.; Visualization, J.C.V. and D.D.G.; Validation, D.D.G.; Data Curation, D.D.G. and M.S.B.; Formal Analysis, M.S.B., D.D.G., J.F.H., R.M.McC., G.A.L., and J.C.V.; Investigation, M.S.B., D.D.G., J.F.H., R.M.McC., G.A.L., and J.C.V.; Writing – Original Draft, J.F.H., D.D.G., and J.C.V.; Writing – Review & Editing, J.F.H., D.D.G., J.C.V., and P.G.F.; Funding Acquisition, P.G.F., J.C.V., and G.A.L.; Resources, J.C.V., P.G.F., and G.A.L.; Project Administration, J.F.H., P.G.F., and J.C.V.; Supervision, J.C.V. and P.F.G.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 6, 2018
Revised: May 23, 2019
Accepted: July 22, 2019
Published: August 8, 2019

## REFERENCES

Altenbach, C., Lopez, C.J., Hideg, K., and Hubbell, W.L. (2015). Exploring structure, dynamics, and topology of nitroxide spin-labeled proteins using continuous-wave electron paramagnetic resonance spectroscopy. Methods Enzymol. *564*, 59–100.

Aziz, A., Hess, J.F., Budamagunta, M.S., FitzGerald, P.G., and Voss, J.C. (2009). Head and rod 1 interactions in vimentin: identification of contact sites, structure, and changes with phosphorylation using site-directed spin labeling and electron paramagnetic resonance. J. Biol. Chem. *284*, 7330–7338.

Aziz, A., Hess, J.F., Budamagunta, M.S., Voss, J.C., and Fitzgerald, P.G. (2010). Site-directed spin labeling and electron paramagnetic resonance determination of vimentin head domain structure. J. Biol. Chem. *285*, 15278–15285.

Aziz, A., Hess, J.F., Budamagunta, M.S., Voss, J.C., Kuzin, A.P., Huang, Y.J., Xiao, R., Montelione, G.T., FitzGerald, P.G., and Hunt, J.F. (2012). The structure of vimentin linker 1 and rod 1B domains characterized by site-directed spin-labeling electron paramagnetic resonance (SDSL-EPR) and X-ray crystallography. J. Biol. Chem. *287*, 28349–28361.

Berendsen, H.J., Postma, J.P., Van Gunsteren, W.F., Di Nola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys. *81*, 3684–3690.

Brown, J.H., Cohen, C., and Parry, D.A. (1996). Heptad breaks in alpha-helical coiled coils: stutters and stammers. Proteins *26*, 134–145.

Case, D.A., Cerutti, D.S., Cheatham, T.E., III, Darden, T.A., Duke, R.E., Giese, T.J., Gohlke, H., Goetz, A.W., Greene, D., Homeyer, N., et al. (2017). AMBER 2017 (University of California, San Francisco).

Chernyatina, A.A., Guzenko, D., and Strelkov, S.V. (2015). Intermediate filament structure: the bottom-up approach. Curr. Opin. Cell Biol. *32*, 65–72.

Chernyatina, A.A., Hess, J.F., Guzenko, D., Voss, J.C., and Strelkov, S.V. (2016). How to study intermediate filaments in atomic detail. Methods Enzymol. *568*, 3–33.

Chernyatina, A.A., Nicolet, S., Aebi, U., Herrmann, H., and Strelkov, S.V. (2012). Atomic structure of the vimentin central alpha-helical domain and its implications for intermediate filament assembly. Proc. Natl. Acad. Sci. U S A *109*, 13620–13625.

Chernyatina, A.A., and Strelkov, S.V. (2012). Stabilization of vimentin coil2 fragment via an engineered disulfide. J. Struct. Biol. *177*, 46–53.

Conway, J.F., and Parry, D.A.D. (1988). Intermediate filament structure: 3. Analysis of sequence homologies. Int. J. Biol. Macromol. *10*, 79–98.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: an N [center-dot] log(N) method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089–10092.

Drozdetskiy, A., Cole, C., Procter, J., and Barton, G.J. (2015). JPred4: a protein secondary structure prediction server. Nucleic Acids Res. *43*, W389–W394.

Edelsbrunner, H., and Koehl, P. (2003). The weighted-volume derivative of a space-filling diagram. Proc. Natl. Acad. Sci. U S A *100*, 2203–2208.

Fiorin, G., Klein, M.L., and Hénin, J. (2013). Using collective variables to drive molecular dynamics simulations. Mol. Phys. *111*, 3345–3362.

Guzenko, D., and Strelkov, S.V. (2018). CCFold: rapid and accurate prediction of coiled-coil structures and application to modelling intermediate filaments. Bioinformatics *34*, 215–222.

Hanukoglu, I., and Fuchs, E. (1983). The cDNA sequence of a Type II cytoskeletal keratin reveals constant and variable structural domains among keratins. Cell *33*, 915–924.

Herrmann, H., and Aebi, U. (2004). INTERMEDIATE FILAMENTS: molecular structure, assembly mechanism, and integration into functionally distinct intracellular scaffolds. Annu. Rev. Biochem. *73*, 749–789.

Herrmann, H., Bar, H., Kreplak, L., Strelkov, S.V., and Aebi, U. (2007). Intermediate filaments: from cell architecture to nanomechanics. Nat. Rev. Mol. Cell Biol. *8*, 562–573.

Herrmann, H., Strelkov, S.V., Feja, B., Rogers, K.R., Brettel, M., Lustig, A., Häner, M., Parry, D.A., Steinert, P.M., Burkhard, P., et al. (2000). The intermediate filament protein consensus motif of helix 2B: its atomic structure and contribution to assembly. J. Mol. Biol. *298*, 817–832.

Hess, J.F., Budamagunta, M.S., Aziz, A., FitzGerald, P.G., and Voss, J.C. (2013). Electron paramagnetic resonance analysis of the vimentin tail domain reveals points of order in a largely disordered region and conformational adaptation upon filament assembly. Protein Sci. *22*, 47–55.

Hess, J.F., Budamagunta, M.S., FitzGerald, P.G., and Voss, J.C. (2005). Characterization of structural changes in vimentin bearing an epidermolysis bullosa simplex-like mutation using site-directed spin labeling and electron paramagnetic resonance. J. Biol. Chem. *280*, 2141–2146.

Hess, J.F., Budamagunta, M.S., Shipman, R.L., FitzGerald, P.G., and Voss, J.C. (2006). Characterization of the linker 2 region in human vimentin using site-directed spin labeling and electron paramagnetic resonance. Biochemistry *45*, 11737–11743.

Hess, J.F., Budamagunta, M.S., Voss, J.C., and Fitzgerald, P.G. (2004). Structural characterization of human vimentin rod 1 and the sequencing of assembly steps in intermediate filament formation in vitro using site-directed spin labeling and electron paramagnetic resonance. J. Biol. Chem. *279*, 44841–44846.

Hess, J.F., Voss, J.C., and FitzGerald, P.G. (2002). Real-time observation of coiled-coil domains and subunit assembly in intermediate filaments. J. Biol. Chem. *277*, 35516–35522.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38, 27-38.

Jeschke, G. (2012). DEER distance measurements on proteins. Annu. Rev. Phys. Chem. *63*, 419–446.

Klare, J.P. (2013). Site-directed spin labeling EPR spectroscopy in protein research. Biol. Chem. *394*, 1281–1300.

Likhtenshtein, G.I. (1993). Biophysical Labeling Methods in Molecular Biology (New York: Cambridge University Press), pp. 46–79.

Liu, L., Hess, J., Sahu, I.D., FitzGerald, P.G., McCarrick, R.M., and Lorigan, G.A. (2016). Probing the local secondary structure of human vimentin with electron spin Echo Envelope modulation (ESEEM) spectroscopy. J. Phys. Chem. B *120*, 12321–12326.

Lupas, A., Bassler, J., and Dunin-Horkawicz, S. (2017). The structure and topology of $\alpha$-helical coiled coils. In Fibrous Proteins: Structures and Mechanisms, D. Parry and J. Squire, eds. (Springer International Publishing), pp. 95–130.

Lupas, A.N., and Gruber, M. (2005). The structure of alpha-helical coiled coils. Adv. Protein Chem. *70*, 37–78.

Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J. Chem. Theory Comput. *11*, 3696–3713.

Meier, M., Padilla, G.P., Herrmann, H., Wedig, T., Hergt, M., Patel, T.R., Stetefeld, J., Aebi, U., and Burkhard, P. (2009). Vimentin coil 1A-A molecular switch involved in the initiation of filament elongation. J. Mol. Biol. *390*, 245–261.

Melo, F., Sanchez, R., and Sali, A. (2002). Statistical potentials for fold assessment. Protein Sci. *11*, 430–448.

Mucke, N., Wedig, T., Burer, A., Marekov, L.N., Steinert, P.M., Langowski, J., Aebi, U., and Herrmann, H. (2004). Molecular and biophysical characterization of assembly-starter units of human vimentin. J. Mol. Biol. *340*, 97–114.

Nagai, K., and Thøgersen, H.C. (1987). Synthesis and sequence-specific proteolysis of hybrid proteins produced in Escherichia coli. Methods Enzymol. *153*, 461–481.

Nicolet, S., Herrmann, H., Aebi, U., and Strelkov, S.V. (2010). Atomic structure of vimentin coil 2. J. Struct. Biol. *170*, 369–376.

Omary, M.B. (2009). "IF-pathies": a broad spectrum of intermediate filament-associated diseases. J. Clin. Invest. *119*, 1756–1762.

Pang, A.H., Obiero, J.M., Kulczyk, A.W., Sviripa, V.M., and Tsodikov, O.V. (2018). A crystal structure of coil 1B of vimentin in the filamentous form provides a model of a high-order assembly of a vimentin filament. FEBS J. *285*, 2888–2899.

Parry, D.A. (2006). Hendecad repeat in segment 2A and linker L2 of intermediate filament chains implies the possibility of a right-handed coiled-coil structure. J. Struct. Biol. *155*, 370–374.

Parry, D.A., and Smith, T.A. (2010). A different conformation for linker L12 in IF molecules in the molecular and filamentous forms: an hypothesis. J. Struct. Biol. *170*, 364–368.

Parry, D.A., and Steinert, P.M. (1999). Intermediate filaments: molecular architecture, assembly, dynamics and polymorphism. Q. Rev. Biophys. *32*, 99–187.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J. Comput. Chem. *26*, 1781–1802.

Polyhach, Y., Bordignon, E., and Jeschke, G. (2011). Rotamer libraries of spin labelled cysteines for protein studies. Phys. Chem. Chem. Phys. *13*, 2356–2366.

Qin, Z., Kreplak, L., and Buehler, M.J. (2009). Hierarchical structure controls nanomechanical properties of vimentin intermediate filaments. PLoS One *4*, e7294.

Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779–815.

Shindyalov, I.N., and Bourne, P.E. (1998). Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Eng. *11*, 739–747.

Sokolova, A.V., Kreplak, L., Wedig, T., Mucke, N., Svergun, D.I., Herrmann, H., Aebi, U., and Strelkov, S.V. (2006). Monitoring intermediate filament assembly by small-angle x-ray scattering reveals the molecular architecture of assembly intermediates. Proc. Natl. Acad. Sci. U S A *103*, 16206–16211.

Steinert, P.M., Jones, J.C., and Goldman, R.D. (1984). Intermediate filaments. J. Cell Biol. *99*, 22s–27s.

Steinert, P.M., Marekov, L.N., and Parry, D.A. (1993). Diversity of intermediate filament structure. Evidence that the alignment of coiled-coil molecules in vimentin is different from that in keratin intermediate filaments. J. Biol. Chem. *268*, 24916–24925.

Steinert, P.M., Steven, A.C., and Roop, D.R. (1985). The molecular biology of intermediate filaments. Cell *42*, 411–420.

Stendardo, E., Pedone, A., Cimino, P., Cristina Menziani, M., Crescenzi, O., and Barone, V. (2010). Extension of the AMBER force-field for the study of large nitroxides in condensed phases: an ab initio parameterization. Phys. Chem. Chem. Phys. *12*, 11697–11709.

Strelkov, S.V., and Burkhard, P. (2002). Analysis of alpha-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. J. Struct. Biol. *137*, 54–64.

Strelkov, S.V., Herrmann, H., Geisler, N., Lustig, A., Ivaninskii, S., Zimbelmann, R., Burkhard, P., and Aebi, U. (2001). Divide-and-conquer crystallographic approach towards an atomic structure of intermediate filaments. J. Mol. Biol. *306*, 773–781.

Strelkov, S.V., Herrmann, H., Geisler, N., Wedig, T., Zimbelmann, R., Aebi, U., and Burkhard, P. (2002). Conserved segments 1A and 2B of the intermediate filament dimer: their atomic structures and role in filament assembly. EMBO J. *21*, 1255–1266.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. *185*, 60–89.

Walshaw, J., and Woolfson, D.N. (2001). SOCKET: a program for identifying and analysing coiled-coil motifs within protein structures. J. Mol. Biol. *307*, 1427–1450.

Wiederstein, M., and Sippl, M.J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. *35*, W407–W410.

# STAR★METHODS

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Bacterial and Virus Strains | | |
| *E. coli* BL21 (AI) | Invitrogen | 1665055EDU |
| pT7 | Invitrogen | 720-0021 |
| Chemicals, Peptides, and Recombinant Proteins | | |
| iProof DNA polymerase | Bio-Rad | 1725301 |
| 1-Oxyl-2,2,5,5-tetramethylpyrroline-3-methyl Methanethiosulfonate (MTSL) | Toronto Research Chemicals | A167900 |
| Source S | GE Healthcare | 17117801 |
| Source Q | GE Healthcare | 17117701 |
| Oligonucleotides | | |
| Oligonucleotides for site-directed mutagenesis (typically 25-35 nucleotides long) were synthesized by Invitrogen. Sequences were identical the vimentin DNA sequence with the exception of the mutated position, at which position the codon was changed to TGC. | Invitrogen | NA |
| Other | | |
| CM-120 electron microscope, biotwin Lens | FEI | Philips CM120 |
| Gatan MegaScan 794/20 digital camera | 2K X 2K | Model 794/20 |
| Gatan Bioscan 792 | Gatan | Model 792 |
| JEOL X-band spectrometer | JEOL-USA | FA-100 and TE100 |
| AKTA FPLC | Amersham Biosciences | UPC-900 |
| Software and Algorithms | | |
| AMBERTOOLS16 | Case et al., 2017 | http://ambermd.org |
| MODELLER | Sali and Blundell, 1993 | https://salilab.org/modeller/ |
| MATLAB | Mathworks, Inc. | https://www.mathworks.com |
| NAMD (CUDA) | Phillips et al., 2005 | https://www.ks.uiuc.edu/Research/namd/ |
| Pymol | Schrodinger, Inc. | https://pymol.org/2/ |
| PROSA | Wiederstein and Sippl, 2007 | https://prosa.services.came.sbg.ac.at/prosa.php |
| TWISTER | Strelkov and Burkhard, 2002 | https://gbiomed.kuleuven.be/english/research/50000715/50000719 |
| UCSF Chimera | Pettersen et al., 2004 | https://www.cgl.ucsf.edu/chimera |
| VMD | Humphrey et al., 1996 | https://www.ks.uiuc.edu/Research/vmd/current/ |
| ALPHASURF | Edelsbrunner and Koehl, 2003 | http://web.cs.ucdavis.edu/~koehl/ |
| Model building and refinement scripts | This paper | jcvoss@ucdavis.edu |
| MD analysis scripts | This paper | jcvoss@ucdavis.edu |
| Deposited Data | | |
| Structure coordinate source | Aziz et al., 2012; | PDB ID: 3UF1 |
| Structure coordinate source | Nicolet et al., 2010 | PDB ID: 3KLT |
| Structure coordinate source | Pang et al., 2018 | PDB ID: 5WHF |
| Structure coordinate source | (Chernyatina and Strelkov, 2012) | PDB ID: 3TRT |
| Structure coordinate source | (Chernyatina et al., 2012) | PDB ID: 3SSU |
| Structure coordinate source | (Chernyatina et al., 2012) | PDB ID: 3SWK |

## LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, John C. Voss (jcvoss@ucdavis.edu).

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

For protein expression, *E. coli* BL21(AI) was used.

## METHOD DETAILS

### Site-Directed Mutagenesis and Purification for Spin Labeling

Vimentin mutants were constructed using site-directed mutagenesis, and recombinant proteins were produced in *E. coli.* Vimentin readily forms inclusion bodies; these are isolated and the recombinant vimentin purified and spin-labeled as described previously (Aziz et al., 2009, 2010; Hess et al., 2002, 2006). In short, site-directed mutagenesis was used to introduce cysteine residues at specific sites in a vimentin expression construct (originally provided by Roy Quinlan, University of Durham, Durham, UK) using Bio-Rad iProof DNA polymerase (Bio-Rad, Hercules, CA) and mutagenic oligonucleotides (Invitrogen, Carlsbad, CA). Coding sequence changes were confirmed by automated DNA sequencing (Davis Sequencing, Davis, CA). Mutant vimentin protein was produced by bacterial overexpression using a pT7 vector (Studier et al., 1990) and *E. coli* BL21AI (Invitrogen, Carlsbad, CA). 1 L of culture in LB medium containing 0.1 mg/ml ampicillin was grown at 37°C to a cell density of 0.6 (according to $OD_{600}$). Thereafter protein expression was induced by adding 1 mM IPTG, followed by 3 additional hr of cultivation at 37°C. Cells were then pelleted by centrifugation and washed in phosphate-buffered saline. Inclusion bodies were purified using lysozyme/DNase, high/low salt washes (Nagai and Thøgersen, 1987), followed by gel filtration and ion-exchange chromatography (AKTA FPLC, GE Healthcare, Piscataway, NJ) as described in (Chernyatina et al., 2016). Site-directed spin labeling was performed by treating the purified protein with 100 μM TCEP (Tris-(2-carboxyethyl-phosphine hydrochloride, Invitrogen) followed by spin labeling with 500 μM thio-specific nitroxide spin label 1-Oxyl-2,2,5,5-tetramethylpyrroline-3-methyl methanethiosulfonate (MTSL; Toronto Research Chemicals, Canada). Unincorporated spin label was removed from spin-labeled vimentin by chromatography over a Source S or Source Q column (both work well (AKTA FPLC, GE Healthcare, Piscataway, NJ)). Protein concentrations were measured by the BCA method (Pierce, Rockford, IL). Purified spin-labeled proteins were stored at -80°C.

### In Vitro Filament Assembly and Electron Microscopy

Filament assembly was conducted by dialyzing the spin labeled protein (~500 μg/ml) overnight in 8 M urea, 50 mM Tris, pH 7.5 against filament assembly buffer (10mM Tris, pH 7.5, 160mM NaCl). Following dialysis, 10 μl of the sample was removed, placed on formvar-coated carbon grids and subsequently stained with 1% uranyl acetate. Grids were examined using a Phillips CM-120 electron microscope, with a Biotwin Lens, (FEI, Hillsboro, OR) operated at 80 kV acceleration voltage. Images were acquired with a Gatan MegaScan 794/20 digital camera (2K X 2K) or a Gatan BioScan 792 (Gatan, Pleasanton, CA).

### EPR Spectroscopy of Site-Directed Spin Labels

EPR measurements of the spin-labeled proteins were conducted on a JEOL X-band spectrometer fitted with a loop-gap resonator (Hess et al., 2002). Spectra were collected from ~6 μl of purified, spin labeled, dialyzed protein, at a final protein concentration of 25–100 μM, loaded in a sealed, quartz capillary tube. Spectra were obtained at room temperature (unless otherwise specified) by a single 2 min scan with 100 Gauss sweep width at a micro-wave power of 4 mW Modulation amplitude (0.125 mT) was optimized to the natural line width of the attached nitroxide as previously described (Aziz et al., 2009, 2010; Hess et al., 2002, 2006). Normalization of the spectra to the same number of spins was done by normalizing each spectrum to the same integrated intensity/amplitude and concentration. To improve the fidelity of the calculation, each sample was double-integrated after solubilization in 2% SDS. Low temperature spectra were collected from samples flash frozen and maintained at -100°C. Protein samples in 5mM Tris, pH 7.5 were mixed with the appropriate volume of 10X IF assembly buffer (100mM Tris, pH 7.5, 1.6M NaCl) in an Eppendorf tube. 25 μl of the mixture was quickly pipetted into a capillary and placed into a low-speed bench top centrifuge to collect the assembling filaments at the bottom. Mixing, pipetting and centrifugation was repeated to generate 2 capillaries for each sample. Both capillaries were placed in the low temperature cavity for low temperature data collection. Distance proximity from frozen spectra was estimated using the semi-empirical broadening parameter $d_1/d$ (Hess et al., 2002; Likhtenshtein, 1993), which was calculated from the intensity of the hyperfine extrema divided by the intensity of the central resonance line.

### Model Building of L1-2 Linker 250-262 (Chain A) and 250-264 (Chain B)

The initial model building of L1-2 linker for chain A and chain B was carried out using BUILD STRUCTURE module of CHIMERA1.11 (Pettersen et al., 2004) using the following input sequence:

$^{230}$EIAFLKKLHEEEIQELQAQIQEQHVQIDVDVSKPDLTAALR$^{270}$

where the phi and psi angle were set at -60, -45 degrees to maintain a right handed alpha-helical geometry with two chains separated by 0.7 nm distance. Next, initial PDB: 3UF1 coordinates of chain A and chain B (Aziz et al., 2012) and PDB: 3KLT coordinates of chain C and chain D (Nicolet et al., 2010) were utilized. For PDB: 3UF1, 34 Selenomethionine (MSE) were mutated to 34 methionine (MET) using a custom python script, while the PDB: 3KLT was unaltered.

### Structural Alignment of 3UFI and 3KLT PDB Coordinates to the L1-2 Linker

PDB: 3UF1 sequences ($^{230}$EIAFLKKLHEEEI$^{242}$) was structurally aligned to L1-2 model via the CEALIGN program (Shindyalov and Bourne, 1998). Next the PDB: 3KLT coordinates were aligned to the C-terminal L1-2 sequence ($^{263}$PDLTAALR$^{270}$) using the rotation and translation function in CHIMERA1.11 (Pettersen et al., 2004). Following this procedure, flanking 230 to 242 model coordinates were deleted and replaced with $^{230}$EIAFLKKLHEEEI$^{242}$ PDB: 3UF1 coordinates, which leads to the L1-2. Similar to findings in the L2 region (Hess et al., 2006), the backbone for sequence positions 263-270 were suggested to be parallel based upon the pattern of EPR-derived $d_1$/d values, and validated upon the linear correlation between $d_1$/d values and the Cα-Cα and the Cβ-Cβ distances of X-ray structural data. Therefore, the L1-2 region was initially set as parallel right-handed helices prior to refinement in the MODELLER program.

### Cα–Cα Spatial Restraints of Chain A to Chain B for Parallel Helix Refinement via MODELLER

In the first step of model refinement, EPR spectra that define $d_1$/d values of 0.42 or greater served as defining positions for the PDB: 3UF1, Linker 1-2 and PDB: 3KLT coordinates. $d_1$/d positions that correspond to the Cα-Cα distance of 0.6 nm between the two chains were defined in the MODELLER (Sali and Blundell, 1993) python script (Table S5). Next, the positions within the region (250-263) for which we do not have the structural Cα-Cα and Cβ−Cβ values were extrapolated from the closest value of X-ray data. For example, the $d_1$/d value for position 250 resembles the values for positions 148, 155, 162, 169, 176, 183, 190, 197, 204, 211, 218, 225, 232, and 243 (having $d_1$/d values of 1.48, 1.35, 1.25, 1.19, 1.19, 1.29, 1.27, 1.11, 1.29, 1.25, 1.25, 1.29, 1.36, and 1.38 nm). Taken together, the average 1.28 nm value serves as a good estimated value for position 250. Subsequently, positions 251 through 263 were extrapolated via the closest previous X-ray structural values and reported in Table S6 and used for the secondary refinement. The third and final refinement was performed using EPR constraints from 124 unique spin-labeled positions corresponding to the structural Cα-Cα and Cβ−Cβ distances, including the extrapolated distances.

The output of the refined model was manually checked for clashes and distances, and the TWISTER (Strelkov and Burkhard, 2002) program was utilized to determine heptad repeats and coiled-coil statistics. TWISTER version 2006 was kindly provided by Professor Sergei V. Stelkov. The full vimentin model was assessed for surface energy using the GA341 energy function (Melo et al., 2002). In addition, the RMSD difference between the models and X-ray structures on the basis of the Z-score was carried out using PROSA (Wiederstein and Sippl, 2007). Finally, vimentin model was assessed using PROSA (Wiederstein and Sippl, 2007) webserver and compared against the MODELLER's energy score function.

Supercoiling of helical strands was evaluated according to the coiled-coil phase yield per residue ($\Delta\omega_n$) of the restrained vimentin model. $\Delta\omega_n$ was calculated using TWISTER and is determined from the average (over both chains) of the mean of dihedral angles formed about the helix and superhelix axes.

### Molecular Dynamics Studies

Molecular dynamics (MD) simulation was performed for the wild-type vimentin and for the 224 MTSL variant vimentin using the hybrid X-ray structures (PDB: 3UF1 and PDB: 3KLT) and EPR restraint L1-2 linker model. All simulations were designed to mimic physiological conditions (e.g. neutral pH). AMBER ff14SB parameters (Maier et al., 2015) were used for amino acids and parm10.dat in AMBERTOOLS16 (Case et al., 2017) was used for cysteine moiety of MTSL and the nitroxide moiety parameters taken from the work of Stendardo et al (Stendardo et al., 2010). MTSL topology was created using the ANTECHAMBER. The model was solvated in a box of TIP3P water (183,879 water molecules in the wild type system and 183,650 water molecules in the 224 MTSL variant system). Appropriate counter ions (Na+ or Cl-) were placed to obtain a neutral system. The system was step-wise minimized. Vimentin was fixed for 3000 steps of Conjugate Gradient (CG) minimization to relax the surrounding water molecules. Next the protein was minimized with 1000 steps of CG, while fixing the water. A final minimization with 1500 steps of CG was carried out to relax the whole system. The CUDA version of NAMD2.12 program (Phillips et al., 2005) was used for the minimization process.

The system was heated incrementally by increasing the temperature from 10K to 298K over 30 picoseconds. It was further equilibrated for 120 picoseconds at a constant pressure of 1 atm using Berendsen coupling method (Berendsen et al., 1984) and constant temperature of 298K using the reassign-hold method (Phillips et al., 2005). During the entire equilibration procedure, the vimentin was restrained with a harmonic potential using a force constant of 2 kcal/mol/Å$^2$ in order to maintain the model. This was done to maintain side chain conformation of the heptad repeat patterns as observed in the X-ray protein crystallography. At the production phase, wild type vimentin restraints were released, temperature was maintained at 298K and 1 atm pressure was held constant using a Berendsen coupling method (Berendsen et al., 1984). In addition, two simulations were carried out using a wild type vimentin model or a 224

MTSL variant model with the addition of harmonic restraint at the *a*,*d* heptad positions using a force constant of 10 kcal/mol/Å$^2$, as well as geometric restraints on the N-terminal and C-terminal ends using a force constant of 0.01 kcal/mol/Å$^2$ and an 0 Å displacement (RMSD and rotation). The restraints were implemented using COLVAR module (Fiorin et al., 2013) integrated in NAMD2.12. Periodic boundary conditions were applied throughout the simulations. Long range electrostatic interactions were treated with particle mesh Ewald [PME] method (Darden et al., 1993) and a switching function was employed between 9 to 12 Å to gradually reduce the electrostatics and van der Waals interactions (Phillips et al., 2005) to zero. The *verlet algorithm* was used for time integration with a time step of one femtosecond. Three independent simulations, each of fifty nanoseconds, were performed for each of the three proteins (wild type without restraints, wild type with restraints, and 224 MSTL variant with restraint). Snapshots were collected every ten picoseconds during these simulations for later analysis.

D257Q, D259G, K262C vimentin mutant variants and a homology model of keratin 14/5 were also created *in silico* using the same procedure described for the vimentin Molecular Dynamics study. Computed Bfactor over the simulations were computed using the following equation:

$$Bfactor(i) = \frac{8}{3}\pi^2 \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(r_i^j - \langle r_i \rangle\right)^2}$$

where $r_i^j$ refers to the position of atom $i$ at timeframe $j$, $\langle r_i \rangle$ is the mean position of the same atom $i$ over the whole simulation, and the summation extends over all $N$ timeframe considered.

The frequency of backbone (N to O) hydrogen bond lost was determined using the following approach. A TCL script is used to produce PDB snapshots from the MD simulation trajectory. Then each PDB snapshot is ran through the find-pairs PYMOL function to generate all backbone hydrogen bonds. Lastly, all backbone hydrogen bonds are filtered using a custom bash script calculated over a 10-residue window with the Hydrogen bond lost frequency computed as $HB_{lost}$ = (H/20) -1. H is identified by the number of actual backbone hydrogen bonds using a cutoff distance of 3.2 Å with an angle of 45° in each snapshot normalized to the 20 possible backbone hydrogen bonds for 10 residues within both chains of the vimentin dimer.

For the above equations, TCL scripts were written and analyzed in the VMD1.9.3 (Humphrey et al., 1996) interface. Visualization and rendering are done using either PYMOL (Schrödinger, LLC) or Blender (http://www.blender.org) program. Plots are generated using MATLAB (The MathWorks, Inc.) platform.

## QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification and statistical analysis of electron paramagnetic resonance data and molecular modeling and the determination of structures by electron paramagnetic resonance are integral parts of existing algorithms and software used which are described in Method Details.

## DATA AND CODE AVAILABILITY

Coordinate files for the vimentin models and scripts written for this study are readily available upon request.