



Disruption of the productive encounter complex results in dysregulation of DIAPH1 activity

Received for publication, August 5, 2023, and in revised form, September 27, 2023. Published, Papers in Press, October 12, 2023.
<https://doi.org/10.1016/j.jbc.2023.105342>

Gregory G. Theophall^{1,‡}, Lisa M. S. Ramirez^{1,‡}, Aaron Premo¹, Sergey Reverdatto¹ , Michaele B. Manigrasso², Gautham Yepuri², David S. Burz¹, Ravichandran Ramasamy², Ann Marie Schmidt², and Alexander Shekhtman^{1,*}

From the ¹Department of Chemistry, State University of New York at Albany, Albany, New York, USA; ²Department of Medicine, Diabetes Research Program, New York University Grossman School of Medicine, New York, New York, USA

Reviewed by members of the JBC Editorial Board. Edited by Phillip A. Cole

The diaphanous-related formin, Diaphanous 1 (DIAPH1), is required for the assembly of Filamentous (F)-actin structures. DIAPH1 is an intracellular effector of the receptor for advanced glycation end products (RAGE) and contributes to RAGE signaling and effects such as increased cell migration upon RAGE stimulation. Mutations in *DIAPH1*, including those in the basic “RRKR” motif of its autoregulatory domain, diaphanous autoinhibitory domain (DAD), are implicated in hearing loss, macrothrombocytopenia, and cardiovascular diseases. The solution structure of the complex between the N-terminal inhibitory domain, DID, and the C-terminal DAD, resolved by NMR spectroscopy shows only transient interactions between DID and the basic motif of DAD, resembling those found in encounter complexes. Cross-linking studies placed the RRRK motif into the negatively charged cavity of DID. Neutralizing the cavity resulted in a 5-fold decrease in the binding affinity and 4-fold decrease in the association rate constant of DAD for DID, indicating that the RRRK interactions with DID form a productive encounter complex. A DIAPH1 mutant containing a neutralized RRRK binding cavity shows excessive colocalization with actin and is unresponsive to RAGE stimulation. This is the first demonstration of a specific alteration of the surfaces responsible for productive encounter complexation with implications for human pathology.

Actin polymerization is a central cellular process that is responsible for cell migration, cytokinesis, and cell proliferation (1, 2). A diaphanous-related formin (DRF), Diaphanous 1 or DIAPH1, is a dimeric, multidomain, multifunctional DRF that possesses actin polymerase activity, promoting the synthesis of linear cables of actin in response to various stimuli (3–6), such as those generated by the signaling of the receptor for advanced glycation end products (RAGE) (7–10). The N-terminal region of DIAPH1 has a GTPase-binding domain, followed by the diaphanous inhibitory domain (DID), that contains an armadillo repeat region (ARR), and an

interdomain helix (IH, also called α 5B), central dimerization (DD), and coiled-coil domains. The C-terminal region of DIAPH1 has the formin homology 1 and 2 domains (FH1 and FH2), which are involved in actin polymerization (3, 4), a T-helix, and the diaphanous autoinhibitory domain (DAD), that binds intramolecularly to DID to maintain DIAPH1 in the autoinhibited state (Fig. 1A) (11–13).

Although no structural information is available for human DIAPH1, structures of the murine homolog of DIAPH1, with 90.3% sequence identity to human DIAPH1 (14), provide a guide to understanding the disparate functions and regulation of DIAPH1. Crystal structures of complexes between the N- and C-terminal fragments of mouse DIAPH1, also called mDia1, present a glimpse into the autoinhibited state (11, 15–17). In DAD, the segment directly N terminal to the RRRK motif forms a helix and a turn, anchored at F1195, that binds to DID to maintain mouse DIAPH1 in an autoinhibited state. The RRRK motif, which is also found in many unrelated proteins, such as coronavirus spike protein (18), proplatelet-derived growth factor A (19), chloroplast signal recognition particle 54 (20), and nuclear valosin-containing protein-like 2 (21), is likely flexible because it is either absent (11, 15) or poorly resolved (16, 17) but is oriented proximal to negatively charged patches on DID (22, 23) and contributes to the binding (11). It is not clear why mutations of this flexible region of DAD result in DIAPH1 activation (22, 24). In addition, the T-helix in mouse DIAPH1 blocks the negatively charged patches on DID, thereby precluding transient interactions between DID and the RRRK motif of DAD (15, 17). Crystal structures of the N- and C-terminal fragments of mouse DIAPH1 display different angles between DID and DD in mouse DIAPH1 dimers, suggesting flexibility in the IH that links DID and DD (11, 15, 23, 25).

DIAPH1 is widely expressed in various cell types, especially muscle (26) and kidney cells (27). Mice devoid of *Diaph1* are viable (27, 28). In humans, loss-of-function mutations in *DIAPH1* are associated with immunodeficiency, microcephaly, and mitochondrial dysfunction (29, 30); whereas mutations in DAD, resulting from deletion or modification of the RRRK motif (22, 31), for example, in the DIAPH1-(R1213X) where the RRRK motif is deleted, confer gain-of-function that have been implicated in autosomal dominant hearing loss (24) and macrothrombocytopenia (32).

[‡] These authors contributed equally to this work.

* For correspondence: Alexander Shekhtman, ashekhtman@albany.edu.

Present address for Lisa M.S. Ramirez: Max Planck Institute for Multidisciplinary Sciences, Goettingen, Germany, 37077.

DIAPH1 and the productive encounter complex

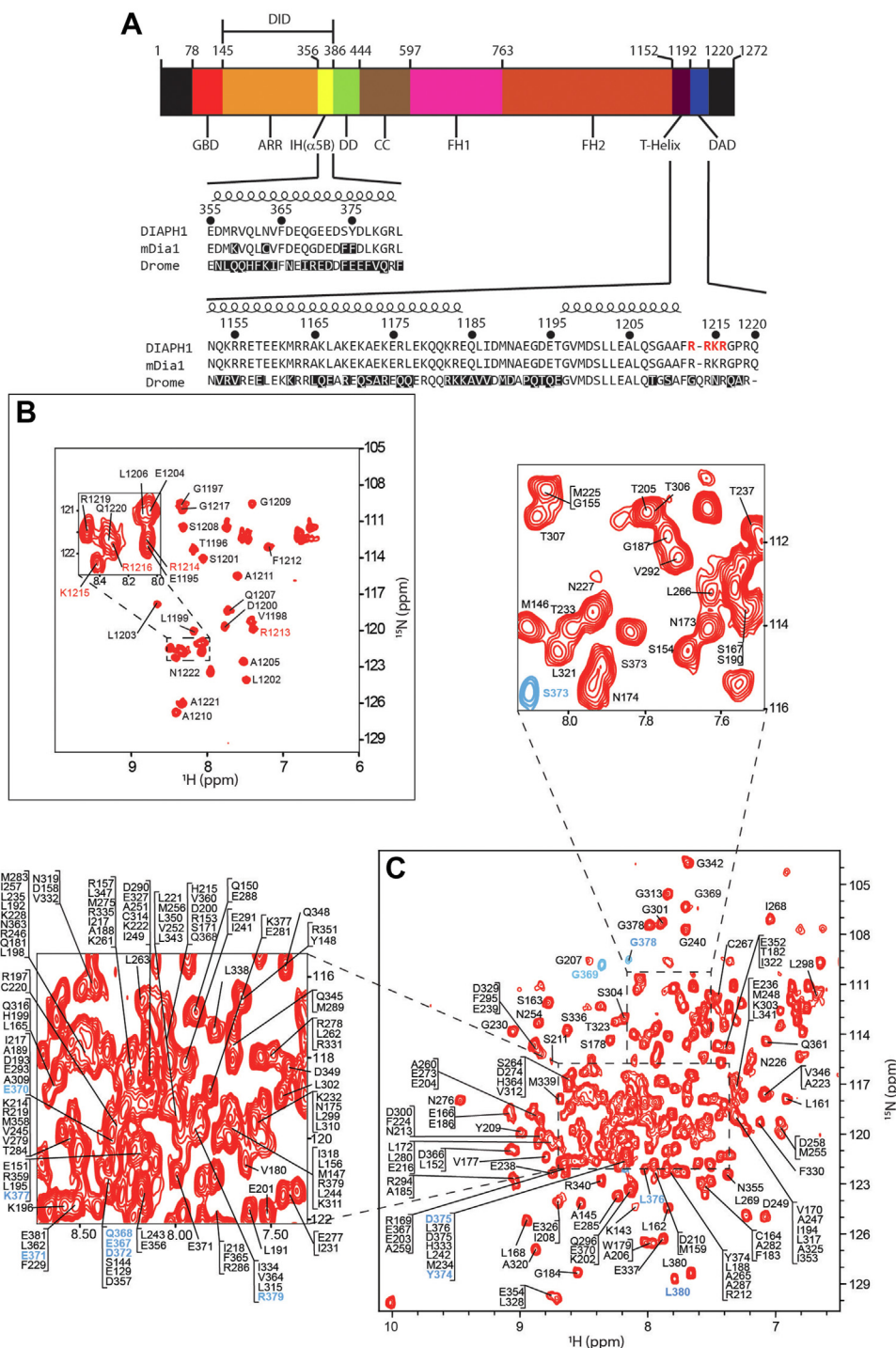


Figure 1. DID and DAD^{M1199L} domains of DIAPH1 form a complex. *A*, schematic representation of the domain structure of DIAPH1. Below are amino acid sequences of the IH and C-terminal region from different organisms: DIAPH1, human (UniProt entry O60610), mouse DIAPH1, mouse (UniProt entry O08808), Drome, fruit fly (UniProt entry P48608). Numbering system corresponds to DIAPH1. Residues that are different from DIAPH1 are highlighted in black, helical regions are indicated by the coiled ribbon above the sequences and RRKR motif is in red. *B*, ¹H-¹⁵N HSQC spectrum of [¹⁵N]-DAD^{M1199L} bound to DID shows well-dispersed backbone amide proton and nitrogen resonances, indicating that DAD^{M1199L} acquires structure upon binding to DID (Fig. S1A). Peaks of the RRKR motif residues are labeled in red. *C*, ¹H-¹⁵N HSQC spectrum of [¹⁵N]-DID bound to DAD^{M1199L} indicates that the complex is structured. Peaks labeled in blue correspond to conformer B. Blue peaks were used in thermodynamic calculations. DAD, diaphanous autoinhibitory domain; DIAPH1, Diaphanous 1; DID, diaphanous inhibitory domain; HSQC, heteronuclear single-quantum coherence.

To understand the structural determinants of DIAPH1 regulation, DID–DAD interactions were characterized by using solution NMR spectroscopy, cross-linking with mass

spectrometry (MS), fluorescence spectroscopy, and functional assays. Flexible regions of DID were identified in the IH. The results also revealed that binding of RRKR to DID is transient

and forms a productive encounter complex (33, 34) that does not compete with the T-helix. These observations allowed a consistent picture of DIAPH1 regulation to be proposed.

Results

DID–DAD^{M1199L} complex stoichiometry and affinity

The DID–DAD interaction was characterized by using a combination of solution NMR spectroscopy and cross-linking MS (Figs. 1 and 2). Residues 142 to 380 of DIAPH1, corresponding to DID, and residues 1194 to 1222, corresponding to DAD, were bacterially overexpressed for labeling and NMR spectroscopy (Fig. 1A). DAD was expressed as an insoluble fusion protein containing an N-terminal his-tag and modified tryptophan leader sequence (35, 36), and the resulting gene product accumulated in inclusion bodies. DAD methionine 1199 was conservatively (37) changed to a leucine, DAD^{M1199L}, to ensure that only a single cyanogen bromide cleavage site was present so that cleavage of the fusion protein would yield

soluble DAD^{M1199L} peptide. This substitution did not change the binding to DID (Figs. S1 and S2).

The stoichiometry of the interaction between DID and DAD^{M1199L} was determined by titrating 100 μM [*L*-¹⁵N]-DAD^{M1199L} with unlabeled DID (Figs. 1B and S1A). Chemical shifts (δ) and cross peak intensities (I) were monitored by using heteronuclear single-quantum coherence (¹H-¹⁵N HSQC) NMR spectroscopy (38) (Fig. S1A). Changes in δ and I were observed at substoichiometric amounts of unlabeled DID. Negligible perturbations occurred beyond a 1:1 mole ratio, indicating a 1:1 stoichiometry for the complex. The strength of the DID–DAD^{M1199L} interaction was measured using tryptophan fluorescence spectroscopy and yielded a macroscopic binding constant of 270 ± 100 nM, in agreement with that reported for WT DAD (11) (Fig. S1B).

DAD^{M1199L} peptide acquires structure when bound to DID

The ¹H-¹⁵N HSQC, spectrum of [*L*-¹⁵N]-DAD^{M1199L}, is shown in Figure 1B; 27 out of 29 residues were assigned.

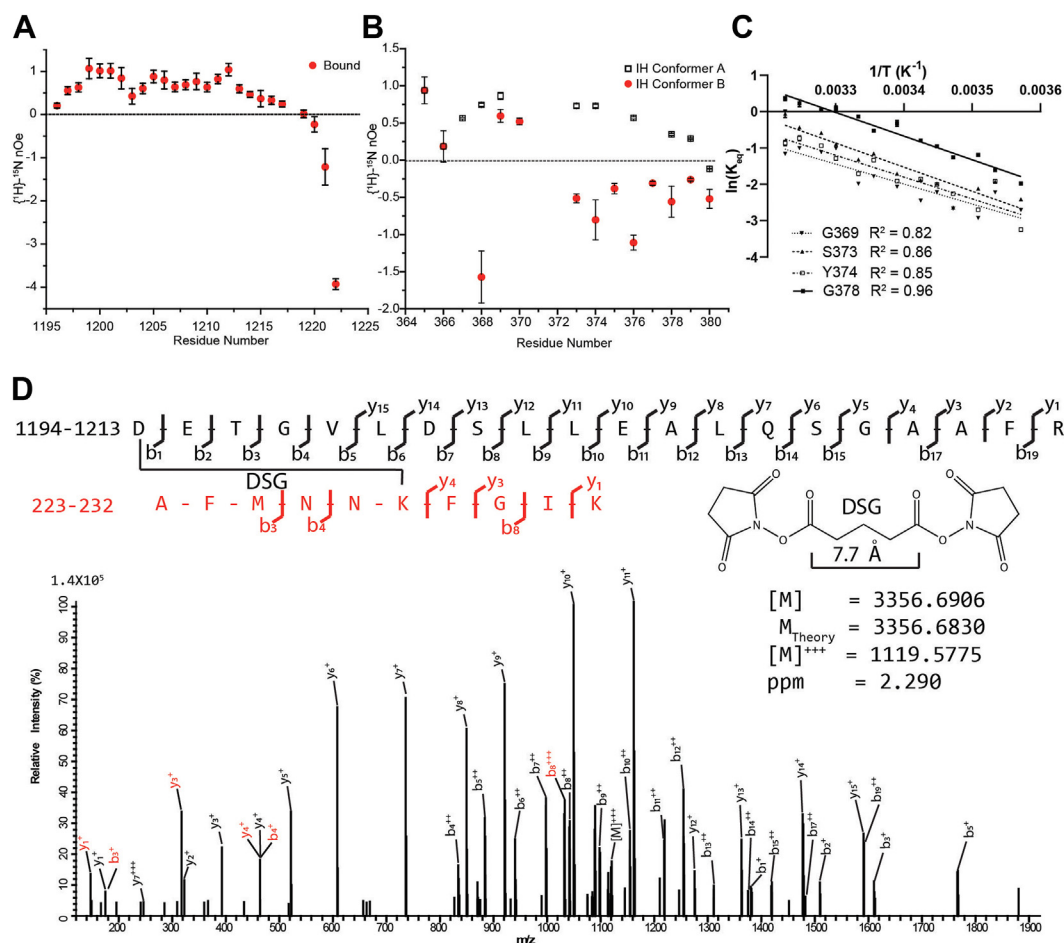


Figure 2. DAD^{M1199L} and interhelical domain of DIAPH1 are flexible. **A**, steady state ¹H-¹⁵N nOes of DAD^{M1199L} bound to DID showed that the structural region of DAD^{M1199L} spans from G1197 to R1213 (Fig. S1A). **B**, steady state ¹H-¹⁵N nOe of the ¹³C-terminal DID residues of IH showed the presence of two conformers: a structured conformer A (white squares) and an unstructured conformer B (red circles). **C**, van't Hoff analysis of selected DID C-terminal residues of IH showed an entropically driven equilibrium between conformers A and B. The temperature dependence of the equilibrium constant, *K*_{eq}, resolved Δ*H*, Δ*S*, and Δ*G* for the unstructured to structured transition (Table S1). **D**, high-energy collision MS spectrum obtained for the DSG cross-linked product between K228 of DID and the N terminus of DAD^{M1199L} at 1119.036 *m/z*. For clarity, only the molecular ion, and the *y* and *b* peptidic fragmentations are labeled. The cross-linked peptides are shown in black and red along with the cross-linker DSG. The experimental mass of 3356.6906 Da is in good agreement with the theoretical mass of 3356.6830 Da calculated from putative elemental composition. Unassigned peaks are from water or ammonia loss, and/or nonpeptidic bond fragmentation. MS data was deposited to <https://repository.jpostdb.org/> under acquisition codes JPST002156 and PXD042130. DAD, diaphanous autoinhibitory domain; DIAPH1, Diaphanous 1; DID, diaphanous inhibitory domain; DSG, disuccinimidyl glutarate.

DIAPH1 and the productive encounter complex

Changes in amide resonances of DAD^{M1199L} upon complex formation with DID suggest that free DAD^{M1199L} undergoes a structural transition when bound (Fig. S1A). Steady-state heteronuclear ¹H-¹⁵N nuclear Overhauser effect (¹H-¹⁵N nOe) (38) was used to identify structured regions in bound DAD^{M1199L} (Fig. 2A). Negative ¹H-¹⁵N nOes indicate a high degree of local flexibility and fast motions on the picosecond to nanosecond time scale (39) consistent with unstructured regions. In contrast, positive ¹H-¹⁵N nOes reflect reduced flexibility indicative of structured regions (39). The ¹H-¹⁵N nOe for the bound DAD^{M1199L} showed that the structured region spans from G1197 to R1213 (¹H-¹⁵N nOes >0.5), while the N and C termini are flexible (Fig. 2A). Importantly, bound DAD^{M1199L} exhibited increasingly negative ¹H-¹⁵N nOes from the positively charged RRRK motif (residues 1213–1216) to the C terminus.

C terminus of DID IH has two conformations

The ¹H-¹⁵N HSQC, spectrum of [*U*-¹⁵N]-DID in the presence of unlabeled DAD^{M1199L}, is shown in Figure 1C. Two hundred twenty-five out of 234 (96%) residues were assigned. Importantly, the M1199L mutation resulted only in minor changes in the chemical shifts of DID, suggesting no major structural changes compared to the WT complex (Fig. S2). Cross peak doubling was observed for 13 C-terminal cross peaks of IH (Fig. 1C blue labels). ¹H-¹⁵N nOes acquired for [*U*-¹⁵N]-DID indicate that the C-terminal residues of IH conformer A exhibit restricted ps-ns dynamics suggesting that conformer A is structured (Fig. 2B). The C-terminal residues of IH conformer B exhibit increased ps-ns dynamics, suggesting that conformer B is unstructured (Fig. 2B). To assess the relative stability of these conformations, van't Hoff analyses of selected C-terminal residues of IH were performed by collecting ¹H-¹⁵N HSQC spectra of [*U*-¹⁵N]-DID bound to DAD^{M1199L} over the temperature range 280 to 310 K. Cross peaks G369, S373, Y374, and G378 were quantified and equilibrium constants for specific residues, K_{eq} , were estimated as the ratio of the amide peak amplitudes of IH conformer A over IH conformer B (Fig. 2C). The negative slopes in Figure 2C showed that the unstructured to structured transition is endothermically disfavored and the y -intercepts in Figure 2C showed that the unstructured to structured transition is entropically favored. When combined, the IH conformer A trends to be thermodynamically disfavored over IH conformer B (Table. S1). The enhanced ps-ns dynamics of IH conformer B compared to conformer A, which results in a decrease in entropy, is likely compensated by the entropically favorable release of bound water, bound ions, or the structural rearrangement of other parts of DID during the B to A transition.

Solution structure of the DID–DAD^{M1199L} complex

The solution structure of a 1:1 complex of DID–DAD^{M1199L}, restricted to IH conformer A of DID, was resolved by using structural restraints derived from solution NMR and cross-linking MS (Figs. 1 and 2D, and Table S3). A standard suite

of NMR experiments (38) was performed to assign backbone and side chain resonances of [*U*-¹⁵N,¹³C]-DID and [*U*-¹⁵N,¹³C]-DAD^{M1199L}. To facilitate unambiguous assignment of DID residues, ¹H-¹⁵N HSQC and 3D ¹H-¹⁵N NOESY (38) spectra of individual ¹⁵N-labeled amino acids of DID bound to WT DAD were recorded at 305 K or 298 K. Intramolecular and intermolecular nOes from 3D ¹H-¹⁵N NOESY spectra were collected for the DID–DAD^{M1199L} complex and converted into upper limits for the proton–proton distances utilized in structural calculations. Distance constraints from MS analyses of DID cross linked to DAD^{M1199L} were obtained by using amine-to-amine chemical cross-linkers. The cross-linked peptides were purified using SDS-PAGE and analyzed using tandem mass spectrometry (MS/MS). Cross-linkers of different lengths gave results for the same set of amines, indicating a flexible and transient structure of the N and C terminus, including the RRRK motif of DIAPH1. (Fig. S3 and Tables S3 and Fig. S4). For this reason, our structure calculations used cross-links that provided tighter distance constraints. A typical spectrum is shown in Figure 2D for the cross-linked product between the N terminus of DAD^{M1199L} and K228 of DID.

Backbone chemical shifts of DID and DAD^{M1199L} were used to predict acceptable ranges of backbone torsion angles per residue using Torsion Angle Likelihood Obtained from Shift and sequence similarity (TALOS)-N software (40). Structural calculations were performed by using Combined assignment and dynamics Algorithm for NMR Applications (CYANA) (41), followed by refinement using Yet Another Scientific Artificial Reality Application (YASARA) (42) in combination with 901 DID intramolecular nOes, 94 DAD^{M1199L} intramolecular nOes, 13 DID–DAD^{M1199L} intermolecular nOes, (Table S2), two DID intramolecular cross-link distances, and two DID–DAD^{M1199L} cross-link distances (Table S3, and Fig. S3). The torsion angles were combined with the distance restraints to yield an ensemble of 20 solution structures for the DID–DAD^{M1199L} complex with a backbone RMSD of 0.3 Å for the helical regions (Fig. 3A), indicating high quality structural convergence. Notably, the loop between α 2A and α 2B, which is absent from the published structures, was well-resolved despite some conformational variability (Fig. 3A, dashed box).

The core of DID presents a five-armadillo repeat arranged in a superhelical coil (Fig. 3B). Despite eleven sequence alterations compared to mouse DIAPH1, P160H, S206A, N208S, Q215K, S307T, E349D, K359R, C363N, D370E, F373S, and F374Y, there were no significant deviations of the ARR from previously published mouse DIAPH1 structures (11, 15–17, 23, 25, 43, 44). Superposition of the ARR in the DIAPH1 DID–DAD^{M1199L} solution structure and mouse DIAPH1 crystal structure (16) yielded a backbone RMSD of \sim 1 Å.

DAD^{M1199L} peptide residues V1198–S1208 formed an amphipathic α -helix that was shorter than previously resolved for mouse DIAPH1 (17, 43) and contacted the concave hydrophobic surface comprised of the central B helices in DID (Fig. 3B). Residues G1209–R1213 turned perpendicular to the helix, anchoring at F1212 (F1195 in mouse DIAPH1), in agreement with published structures (11, 15–17), while the

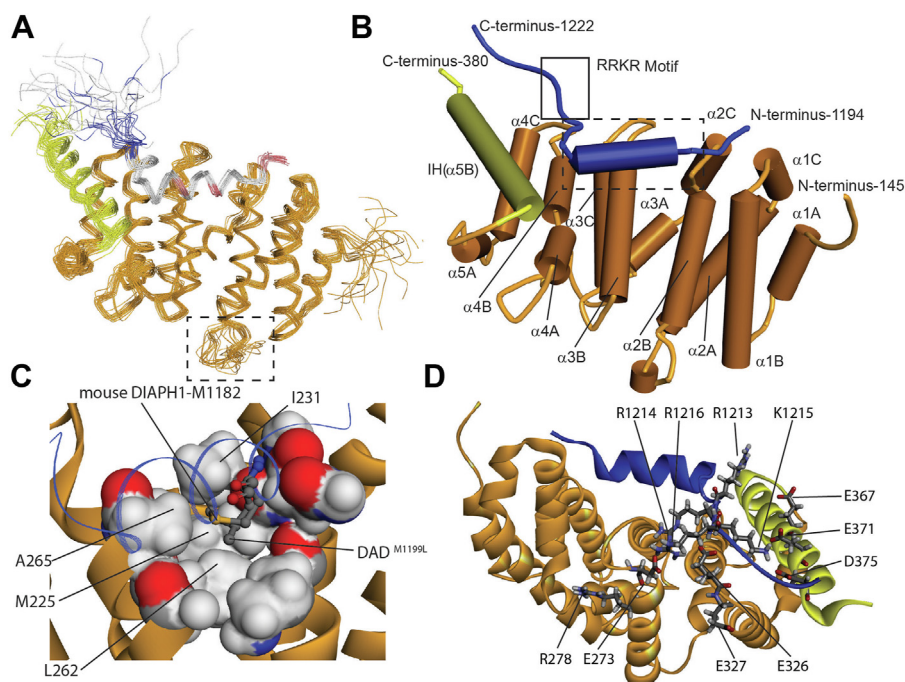


Figure 3. Solution structure of human DIAPH1 DID-DAD^{M1199L} complex reveals the binding mode of the RRKR motif. *A*, overlay of the backbone traces of the 20 solved DID-DAD^{M1199L} solution structures (PDB code 8FG1) refined using YASARA (42). The ARR region of DID is in orange, the IH is in yellow, and the DAD^{M1199L} peptide is colored by pKa where blue is basic, gray is neutral, and red is acidic. The dashed box outlines residues absent from published structures. *B*, tube diagram of a representative DID-DAD^{M1199L} structure showing the superhelical orientation of the armadillo repeats of DID and the DAD^{M1199L} peptide (blue). The dashed box outlines in panel C. The solid box outlines the RRKR motif in panel D. *C*, close up of the DID van der Waals surface of residues within 5 Å of L1199. White is hydrogen, red is oxygen, blue is nitrogen, gray is carbon, and orange is sulfur. M1182 of mouse DIAPH1, which corresponds to M1199 of DIAPH1, and L1199 side chains are shown as balls and sticks. The DAD^{M1199L} peptide backbone (this work) is shown as a thin blue ribbon. M1182 is from PDB 2F31. *D*, DID-DAD^{M1199L} complex, rotated 180° horizontally and downward 90° relative to 3B, showing putative interactions between the basic RRKR motif and acidic residues of DID. Residues E273, R278, E367, E371, and D375 were subjects of previous mutagenesis studies (23, 25). All models were created in Discovery Studio. ARR, armadillo repeat region; DAD, diaphanous autoinhibitory domain; DIAPH1, Diaphanous 1; DID, diaphanous inhibitory domain; IH, interdomain helix.

positively charged C terminus remained unstructured. The L1199 side chain had a dihedral angle of $\chi_1 = -160^\circ$, comparable to the -177° observed for the corresponding residue, M1182, in the crystal structure of the complex formed by the C- and N-terminal fragments of mouse DIAPH1 (16). L1199 also has similar χ_2 and χ_3 values and resides in a hydrophobic pocket contacting M225, I231, L262, and A265, in agreement with published structures (Fig. 3C) (11, 15–17). The results confirmed the NMR observations (Fig. S2) and demonstrated that the leucine substitution does not perturb the WT structure.

Encounter complexes are reaction intermediates that arise from an ensemble of conformations that can sample various orientations to properly align binding determinants (33, 34, 45). As postulated for other cases of encounter complexes (33, 34, 45), DAD residues R1213–R1216 are electrostatically steered (34, 45) toward the negatively charged surface of DID. The RRKR motif does not exhibit NMR-visible interactions with DID and the entirety of this highly basic region is absent in the published structures. Cross-linking together with MS/MS experiments showed that DAD^{M1199L} K1215 is within 11.4 Å of DID K228 and 11.4 Å of DID K377 (Table S3). In addition, MS analysis yielded cross-links for the K228–K1215 pair obtained using a 21.7 Å linker (Table S4). Both DID lysine residues are too far apart to interact simultaneously with DAD^{M1199L} K1215, indicating a high degree of flexibility for

the C terminus of DAD^{M1199L} in agreement with the NMR data. When either the DAD^{M1199L} K1215/DID K377 or DAD^{M1199L} K1215/DID K228 cross-link was used for structure calculations, the RRKR motif interacted with the acidic DID residues E326, E367, E371, and D375, in agreement with mutagenesis studies (23, 25) (Figs. S4 and 3D). These residues constitute a negative patch on DID that interacts with the basic RRKR motif of DAD^{M1199L}, suggesting that electrostatic steering may play a role in complex formation as previously described for productive encounter complexes (33, 34, 45, 46).

T-helix does not impede DID–DAD^{M1199L} interactions

Available structures of DRFs revealed a long helical structure preceding DAD^{M1199L}, the T-helix, which also contacts DID (15, 17). This helix was poorly resolved in the crystal structure suggesting flexibility (15, 17). We used NMR spectroscopy to resolve this element and observed a well-dispersed ¹H-¹⁵N HSQC spectrum of the T-helix, which indicated the presence of a folded structure (Fig. S5A and Table S5). Analyses of amide–amide cross-peaks in the ¹H-¹H NOESY spectrum (Fig. S5B) and chemical shift indices (Fig. S5C) revealed a transient helical structure in solution. Superimposing the DID-DAD^{M1199L} structure from this work with the T-helix and DAD from mouse DIAPH1 (15) shows that the RRKR motif of DAD^{M1199L} appears to sterically clash with the

DIAPH1 and the productive encounter complex

T-helix (Fig. S5D). Such an occlusion could prevent the formation of an encounter complex by masking the acidic patch on DID. Contribution of the RRKR motif to the overall binding affinity of DID for DAD^{M1199L} can be estimated from binding studies¹¹ to be in the millimolar range, assuming that binding of the DAD helix and the RRKR motif to DID are cooperative. To probe this interaction, 120 μM of [U -¹⁵N]-DAD^{M1199L} bound to unlabeled DID was titrated with 360 μM of purified T-helix and analyzed *via* NMR (Fig. S5E). The absence of chemical shift perturbations in [U -¹⁵N]-DAD^{M1199L} indicates that there are no T-helix residues that compete with RRKR for binding to DID.

Double mutant DID^{E326G/E327A} alters the affinity and kinetics of DAD^{M1199L} binding

To biochemically test if the acidic patch on DID (Fig. 3D) is involved in electrostatic steering (34, 45), mutations E326G/E327A, DID^{E326G/E327A}, were introduced. These mutations change the electrostatic character of the patch converting it from negatively charged to mostly neutral (Fig. 4A). The ¹H-¹⁵N HSQC spectrum of [U -¹⁵N]-DID^{E326G/E327A} complexed with DAD^{M1199L} exhibited a well-dispersed amide envelope and minor chemical shift changes (Fig. S6, A and B) as compared to that of the WT [U -¹⁵N]-DID-DAD^{M1199L} complex, indicating that the tertiary structure of DID^{E326G/E327A} is comparable to that of the WT.

To test if the DID negative patch is critical for DAD binding, tryptophan fluorescence titrations were performed with 200 nM DID^{E326G/E327A} and DAD^{M1199L} concentrations ranging from 10 to 90 μM (Fig. S1C). The K_D resolved for DID^{E326G/E327A} was $1.5 \pm 0.5 \mu\text{M}$, which is about five times higher than that for WT DID (Fig. S1B). This result suggests that the interaction between the RRKR motif of DAD^{M1199L}

and the negatively charged patch of DID is specific and increases the overall affinity of DAD^{M1199L} for DID.

Stopped flow fluorescence titrations were used to evaluate the effect of the E326G/E327A mutation on the DID-DAD^{M1199L} association kinetics. Two hundred nanomolar (200 nM) DID and DID^{E326G/E327A} were titrated with 1 to 20 μM DAD^{M1199L} to create pseudo first order kinetics. The data were fit to a single exponential to estimate the observable rate constants, k_{obs} (Fig. 4B). A k_{on} of $52 \pm 1.5 (\mu\text{M s})^{-1}$ and a k_{off} of $19 \pm 17 \text{ s}^{-1}$ were resolved for the DID-DAD^{M1199L} interaction and a k_{on} of $13 \pm 1 (\mu\text{M s})^{-1}$ and a k_{off} of $27 \pm 14 \text{ s}^{-1}$ for the DID^{E326G/E327A}-DAD^{M1199L} interaction. The 4-fold decrease in k_{on} suggests that the decreased binding affinity of DAD^{M1199L} to DID^{E326G/E327A} is primarily due to a decrease in the rate of association, in agreement with an electrostatic steering mechanism.

DAD^{M1199L} also forms futile encounter complexes by nonspecific binding to DID

The model used to fit the tryptophan fluorescence data (Fig. S1) could not discriminate between specific and nonspecific interactions. To identify residues involved in nonspecific interactions, [U -¹⁵N]-DID was titrated with DAD^{M1199L} and changes in the ¹H-¹⁵N HSQC spectra were monitored (Fig. 5A). With a large molar excess of DAD^{M1199L}, nonspecific binding interactions could be inferred from changes in DID chemical shifts. We observed minor chemical shift perturbations of DID residues (Fig. 5B), most of which were clustered in at least four noncontiguous areas on the DID surface (Fig. 5C). Thus, our chemical shift mapping reveals nonspecific binding sites for DAD^{M1199L} that could give rise to so-called futile encounter complexes (34).

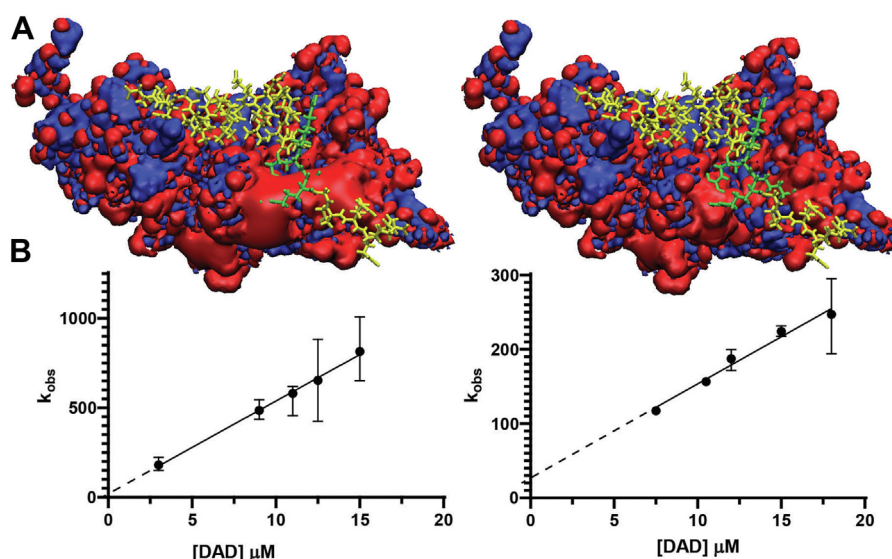


Figure 4. The acidic patch of DID is required for strong DAD^{M1199L} binding. A, iso-potential surface of DID (left) and DID^{E326G/E327A} (right) bound to DAD^{M1199L} where blue and red represents +10 eV and -10 eV, respectively. The basic RRKR motif of DAD^{M1199L} is in green and the rest of DAD^{M1199L} is in yellow. The E326G and E327A mutations reduce a large negatively charged surface on DID as evidenced by the relocation of the basic side chain (green) above the -10 eV iso-contour surface (right). The orientation is analogous to Figure 3D. B, stopped flow fluorescence titrations for the association kinetics of DID-DAD^{M1199L} (left) and DID^{E326G/E327A}-DAD^{M1199L} (right). Surface models of DID and residues of DAD^{M1199L} were made by using Visual Molecular Dynamics software (65). All graphs were made using GraphPad Prism (graphpad.com). DAD, diaphanous autoinhibitory domain; DID, diaphanous inhibitory domain.

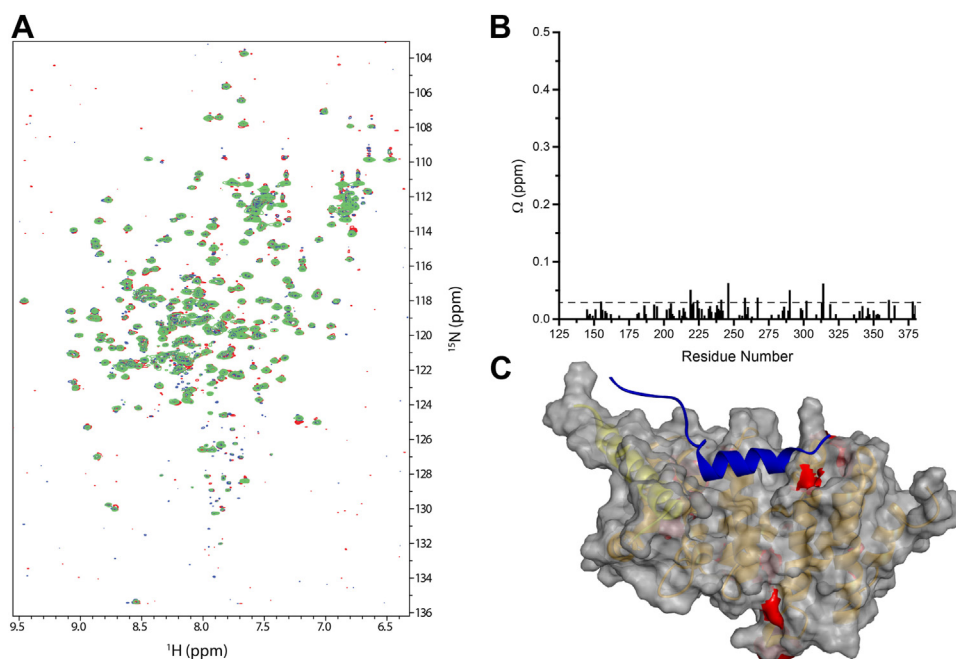


Figure 5. Nonspecific binding of DAD^{M1199L} results in the formation of futile encounter complexes. *A*, superimposed ¹H-¹⁵N-HSQC spectra of 100 μM [^U-¹⁵N] DID in the presence of a 1× (*red*), 10× (*blue*), and 18× (*green*) molar excess of DAD^{M1199L}. *B*, chemical shift changes, Ω, in the ¹H-¹⁵N-HSQC spectra of [^U-¹⁵N]-DID in the presence of a 1× and 18× molar excess of DAD^{M1199L}. $\Omega = \sqrt{(\Delta H)^2 + (\Delta N/2)^2}$, where ΔH and ΔN are changes in chemical shift in the hydrogen and nitrogen dimensions, respectively. *C*, semitransparent space filling model of DID showing significant chemical shift changes (*red*) upon nonspecific DAD^{M1199L} binding. The cut-off value for significant changes, Ω = 0.03, defines the largest 5% of the Ω values. The core structure of DID is *tinted brown* and the interdomain helix is *yellow*. The orientation of the DAD^{M1199L} peptide in the encounter complex is shown in *blue*. DAD, diaphanous autoinhibitory domain; DID, diaphanous inhibitory domain; HSQC, heteronuclear single-quantum coherence.

Mutations in the negative patch of DID result in dysregulation of DIAPH1 activity

Autoinhibited DIAPH1 is free and cytosolic; however, upon activation it binds to the ends of actin fibers (3, 4). Deletion of the DAD sequence results in activation of DIAPH1 and increased actin binding (11, 22, 23). Importantly, deletion of the RRKR motif in DIAPH1 also results in elongation of microvilli with subsequent increased DIAPH1-actin colocalization (24). We hypothesized that mutations E326G/E327A in DID, which disrupt the binding of RRKR and thus the formation of the productive encounter complex, would also lead to increased DIAPH1-actin colocalization.

To test this hypothesis, the DAD deletion mutant (12) (DIAPH1^{ΔDAD}), which was used as a positive control, and double mutant DIAPH1^{E326G/E327A}, were introduced into HEK293T cells. All constructs were expressed with an N-terminal GFP tag to monitor the location of the proteins by fluorescent microscopy. F-actin was tagged with phalloidin-Alexa Fluor 568 to measure the colocalization of DIAPH1-GFP constructs with F-actin (Fig. 6A). Manders coefficients, M1 and M2 (47), were used to quantify the colocalization of F-actin with DIAPH1-GFP constructs and DIAPH1-GFP constructs with F-actin, respectively (Fig. 6B). As expected, both DIAPH1^{ΔDAD}-GFP and DIAPH1^{E326G/E327A}-GFP exhibited significant increases in colocalization with F-actin compared to that of the WT.

To further test the functional importance of the RRKR motif, we examined RAGE signal transduction (7–9). DIAPH1 is an intracellular effector of RAGE; stimulation of RAGE by

RAGE ligand AGEs, such as carboxymethyllysine human serum albumin (CML-HSA) results in a DIAPH1-dependent increase in migration of vascular smooth muscle cells (SMCs) (7–9). Human SMCs were transfected with plasmids-expressing yellow fluorescent protein (YFP), DIAPH1-YFP, DIAPH1^{ΔDAD}-YFP, and DIAPH1^{E326G/E327A}-YFP, and the cells were treated with 1 μM of CML-HSA (Fig. 6C). SMCs transfected with DIAPH1-YFP exhibited an insignificant increase in migration upon stimulation compared to those transfected with YFP, suggesting that endogenous DIAPH1, which is present in SMCs, is sufficient to support RAGE-dependent cell migration. SMCs transfected with the positive control DIAPH1^{ΔDAD}-YFP exhibited a statistically significant reduction in cell migration upon stimulation, suggesting that the deletion of DAD disrupts RAGE signal transduction. SMCs transfected with DIAPH1^{E326G/E327A}-YFP also exhibited a statistically significant reduction of cell migration upon stimulation, although this reduction trended lower than that of the positive control.

Discussion

The solution structure of human DID-DAD^{M1199L} elucidated features of autoinhibited human DIAPH1 that were unresolved in previous crystal structures of the mouse analog to reveal important aspects of DIAPH1 dynamics and its role in regulation. Part of the IH of DIAPH1, residues 367 to 380, exhibits an equilibrium between structured and unstructured conformers. The structured conformation is helical and similar to that detected in crystalline mouse DIAPH1 DID (11, 15–17,

DIAPH1 and the productive encounter complex

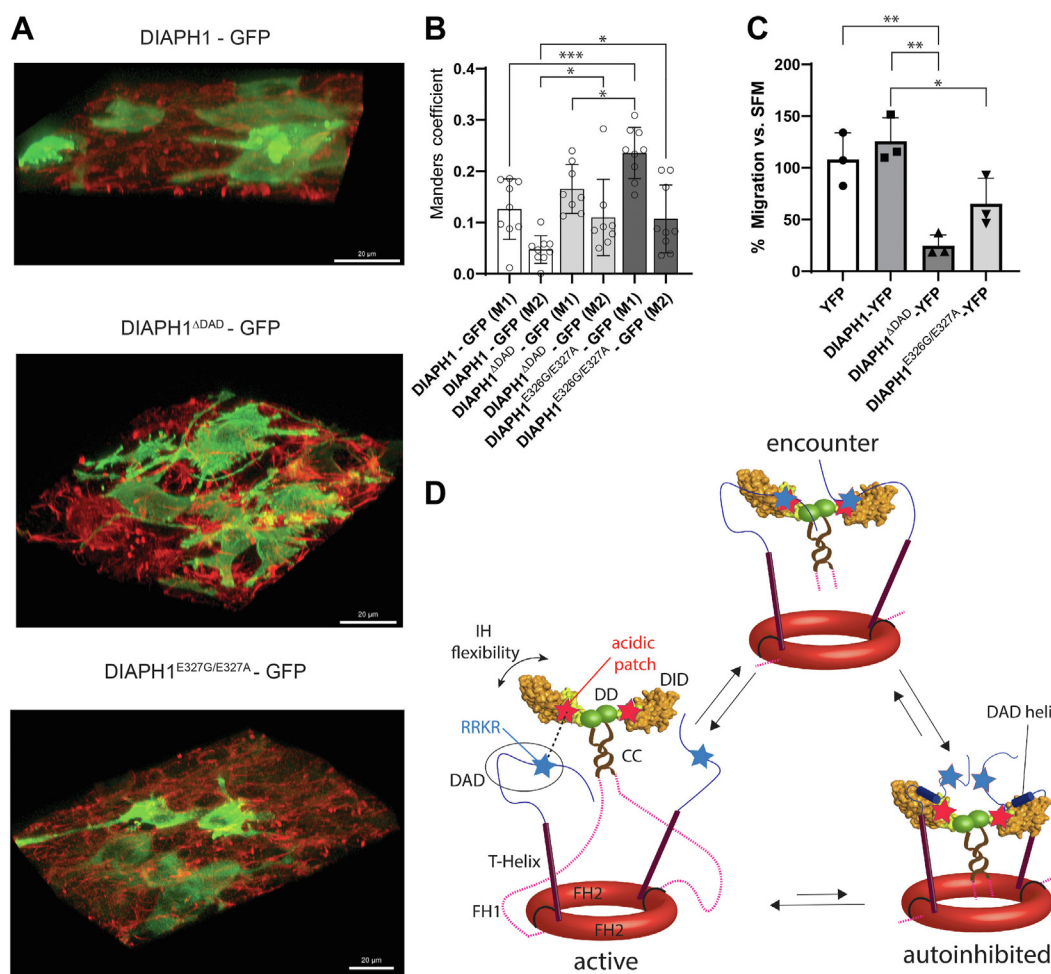


Figure 6. Formation of a productive encounter complex increases DIAPH1-actin colocalization and regulates RAGE ligand-induced cellular migration in human vascular smooth muscle cells. *A*, fluorescence microscopy of HEK293T cells with GFP-tagged WT DIAPH1 (*top*), DIAPH1^{ΔDAD} (*middle*), and DIAPH1^{E326G/E327A} (*bottom*) in *green*, and actin labeled with Alexa Fluor 568 Phalloidin in *red*. *Yellow* indicates colocalization. *B*, manders coefficients, M1 and M2, quantifying colocalization between F-actin and DIAPH1 constructs and DIAPH1 constructs and F-actin, respectively. Larger values correspond to a higher degree of colocalization. Column heights reported on the graph represent mean values, error bars represent the SD. *C*, percent migration of human primary aortic vascular smooth muscle cells, SMCs, when stimulated with CML-HSA as compared to those treated with serum-free medium (SFM). Cells expressing YFP and DIAPH1-YFP were used as a reference and compared with those expressing DIAPH1^{ΔDAD}-YFP and DIAPH1^{E326G/E327A}-YFP. *Stars* indicate statistical significance: *p*-value < 0.05 (*), *p*-value < 0.01 (**), and *p*-value < 0.001 (***). *D*, proposed model for the equilibrium of WT DIAPH1 dimers between the active, encounter, and autoinhibited complexes. *Double headed arrow* indicates IH flexibility, and unstructured DAD is shown in the *circle*. Colored domains match that of [Figure 1A](#). In mouse DIAPH1, DD, CC, and FH2 domains facilitate dimerization ([15](#), [17](#), [43](#), [72](#)). Active DIAPH1 forms a productive encounter complex *via* electrostatic steering, *dotted line*, between the RRKR motif (*blue star*) and the acidic patch of DID (*red star*). The interaction increases the probability of DID-DAD helix coupled folding and binding and formation of the autoinhibited conformation. Active DIAPH1 lacking the basic RRKR motif, such as in DIAPH1-(R1213X) ([22](#), [31](#)), cannot create a productive encounter complex resulting in an equilibrium shift toward the active complex. Note that deletion of DAD, such as in DIAPH1^{ΔDAD}, would preclude the autoinhibition. CC, coiled-coil; CML-HAS, carboxymethyllysine human serum albumin; DAD, diaphanous autoinhibitory domain; DIAPH1, Diaphanous 1; DID, diaphanous inhibitory domain; IH, interdomain helix; RAGE, receptor for advanced glycation end products; YFP, yellow fluorescent protein.

[23](#), [25](#), [43](#), [44](#)). Thermodynamic data show that the structured form is likely favored but still retains enough flexibility observed in variability between DID and DD domains in crystalline forms of mouse DIAPH1 ([11](#), [15](#), [23](#), [25](#)). Flexibility in the regions between the DID and the rest of DIAPH1 contributes to the formation of the autoinhibited state by increasing the probability of DID, encountering DAD in the context of full-length DIAPH1.

The lack of structural information on the DAD C terminus complicates efforts to understand atomic-level interactions between the basic RRKR motif and DID implicated in human pathologies ([24](#), [29](#), [30](#), [32](#)). Previous studies have shown that the RRKR motif is not essential for DID-DAD binding but

does contribute to the binding affinity ([11](#), [22](#), [31](#)). Removal of the RRKR motif decreases DID-DAD binding affinity by 5-fold ([11](#)), resulting in gain-of-function ([24](#)). Our NMR and cross-linking studies show that the RRKR motif is flexible in the DID-DAD^{M1199L} complex ([Tables S3](#) and [S4](#)). How could such a strong effector of K_D be flexible? Findings from cross-linking and stopped flow experiments in this work reveal transient, yet specific, interactions involving RRKR, thus providing insight into a possible structural mechanism of autoinhibitory regulation.

Earlier structural studies revealed that Rho GTPase binding to DIAPH1 elicits a conformational change that displaces the DAD helix from DID ([12](#), [25](#), [48](#)). These studies showed that

the binding causes only partial activation of DIAPH1, suggesting that other regulatory steps are involved in DIAPH1 activity. Our study further elaborates on this point by at least partially implicating the formation of an endogenous encounter complex in the regulation of DIAPH1 activity.

Previous mutagenesis studies showed that mutating DID residues E367R, E371R, or D375R (23), which are located in the vicinity of the RRKR-binding site (Figs. 3D and S4), resulted in a 2- to 3-fold decrease in DID-DAD binding affinity, but E273K and R278E did not (25). Our findings identify E326 and E327 as critical for binding (Figs. 3D, 4 and S1). The E326 and E327 acidic patch is located greater than 30 Å away from the Rho GTPase binding site and represents a new surface for DIAPH1 regulation (Fig. S4B) (25). Interruption of the electrostatic interaction between the basic RRKR motif and these acidic residues decreased DID-DAD^{M1199L} binding affinity by a factor of five. Due to the long-range nature of electrostatic interactions, the combined effect of E326G/E327A mutations with those previously studied (23) is expected to be cumulative. Importantly, the E326G/E327A mutations that disrupt the RRKR binding surface also increased colocalization of actin and DIAPH1 *in vivo* (Fig. 6, A and B), consistent with previous observations that deletion of RRKR leads to DIAPH1 colocalization with actin in filopodia (24).

DIAPH1 is critically involved in RAGE signaling: SMC migration is increased when RAGE is activated by its ligand CML-HSA and this increase depends on RAGE-DIAPH1 signal transduction (8–10). Perhaps counterintuitively, the E326G/E327A mutation resulted in significant reduction of SMC migration. Note that efficient cellular migration relies on regulation of actin polymerization/depolymerization and, thus, dysregulation leads to reduced migration (49–51). This suggests that DIAPH1^{E326G/E327A} was not fully regulated by autoinhibition (Fig. 6C). An even larger reduction was observed in the SMCs expressing DIAPH1^{ADAD} where the regulatory DID-DAD interaction is completely removed; DIAPH1^{ADAD} was used as a positive control for dysregulated constitutively active DIAPH1.

Because the K_D for DID-DAD^{M1199L} increased when DID E326G/E327A mutations were introduced and the interaction remained transient, it is likely that DID^{E326G/E327A} mutations affect primarily the association of DAD^{M1199L} to DID. Kinetic experiments and NMR titrations suggest that RRKR interacts with the acidic residues on DID and this interaction is not obstructed by the T-helix (15, 17). Altogether, these observations paint a dynamic picture. The conformational exchange between structured and unstructured in the IH segment of DID allows the DID to search more conformational space. This, in turn, allows the basic RRKR motif to interact with the acidic patch of DID facilitating the wobbling and searching (52), folding and binding (53) of the DAD helix. This interaction is absent in RRKR deletion mutants leading to shifting equilibrium to the active form of DIAPH1 (Fig. 6D) (22, 31). Subsequent exchange of RRKR with the basic portion of the T-helix may then free the C terminus of DAD. This agrees with the current understanding that the RRKR motif interacts with actin and facilitates actin fiber nucleation (22, 54).

Importantly, the mechanism proposed here describes a productive encounter complex between DID and DAD required for autoinhibition, which can be disrupted by mutations in DID or deletions in DAD leading to gain-of-function activity. This suggests that the disease states characterized by gain of function, such as autosomal dominant hearing loss (24) and macrothrombocytopenia (32), may arise from enhanced actin assembly mediated by DIAPH1 dysregulation.

Experimental procedures

DID cloning, expression, and purification

Expression plasmid pET-28a(+)-DID, which confers kanamycin resistance and expresses an N-terminal his-tagged DID, residues 142 to 380 from human DIAPH1 (UniProt O60610), was purchased from Genscript. pET28a(+)-DID was transformed into *Escherichia coli* strain BL21(DE3) and grown overnight at 37 °C in 100 ml of LB containing 35 µg/ml of kanamycin (LB-Kn). The culture was transferred to 1 L of LB-Kn and incubated at 37 °C until the A_{550} reached 0.90. To overexpress DID without stable isotope labeling, IPTG was added to a final concentration of 1 mM and the cells were incubated for 4 h at 37 °C. For isotopic labeling, cells were harvested by centrifugation at 4000g and resuspended in minimal (M9) medium, 5 mM Na₂HPO₄, 2.5 mM KH₂PO₄, pH 7, 1 mM NaCl, 2 mM MgSO₄, 0.1 mM CaCl₂, and 1 mg/l thiamine hydrochloride, supplemented with 35 µg/ml of kanamycin (M9-Kn). For [¹⁵N] labeling, 1 g/l of ¹⁵N-ammonium chloride (Sigma-Aldrich) was used as the sole nitrogen source. For [¹⁵N,¹³C] labeling 2 g/l of ¹³C-glucose (Sigma-Aldrich) was used as the sole carbon source. Overexpression of labeled DID was induced with 1 mM IPTG and proceeded for 4 h at 37 °C. Cells were centrifuged at 4000g, resuspended in lysis buffer, 20 mM Hepes, pH 8, 1 M NaCl, 12.5% w/v sucrose, 4 M urea, and 10 mM β-mercaptoethanol, containing an EDTA-free protease inhibitor tablet (Roche) and sonicated with a Model 250 Digital Sonifier (Branson) at 40% amplitude using 2.5-min cycles of 0.3 s pulses interrupted by 1 s pauses. The lysate was centrifuged for 40 min at 20,000g and 4 °C. The supernatant was loaded into a nickel-nitriloacetic acid agarose affinity column (Qiagen) pre-equilibrated with lysis buffer. The partially denatured protein bound to the column was refolded by washing with five column volumes of refolding buffer, 10 mM Hepes, pH 8, 300 mM NaCl, and 10 mM imidazole. The column was washed with five column volumes of wash buffer, 10 mM Hepes pH 8, 300 mM NaCl, and 20 mM imidazole, and the refolded DID was eluted with elution buffer, 10 mM Hepes, pH 8, 300 mM NaCl, and 250 mM imidazole. EDTA was added from a 0.5 M pH 8 stock to a final concentration of 10 mM, and the eluant was dialyzed into ion exchange chromatography buffer A, 20 mM NaH₂PO₄/Na₂HPO₄, pH 8.10, 100 mM NaCl, prior to loading onto a Hi-Trap Q-XL anion exchange column (Cytiva). A linear gradient of buffer B, 20 mM sodium phosphate, pH 8.10, 1 M NaCl, was used to elute the protein at a conductivity of ~40 mS/cm. The purity of the DID protein (molecular weight of 29,078 Da) was found to be >95% on a

DIAPH1 and the productive encounter complex

12% SDS-PAGE, and the concentration was determined by absorbance using a molar extinction coefficient of $10,220 \text{ M}^{-1} \text{ cm}^{-1}$ at 280 nm.

DID^{E326G/E327A} was overexpressed using the plasmid pET-28a(+)-DID^{E326G/E327A} (GenScript), which confers kanamycin resistance and expresses an N-terminal his-tagged, mutated DID^{E326G/E327A}. Purification was performed using the protocol identical to that of DID.

Residue-selective labeling of DID was accomplished by transferring the bacterial cells grown in 1 L of LB-Kn to 1 L of M9-Kn medium containing a mixture of 19 unlabeled amino acids plus 0.5 g/l of the ¹⁵N-labeled amino acid of interest (55). Overexpression was carried out for 4 h at 21 °C after adding IPTG to 1 mM. The cells were harvested and selectively labeled DID protein was purified using the methodology described above for [¹⁵N]-labeled DID.

DAD and DAD^{M1199L} cloning, expression, and purification

The 29-residue DAD peptide, DETGVMSLLEALQSGAARFRKRGRQAN, corresponding to WT residues 1194 to 1222 (UniProt O60610), was purchased from GenScript. Stock solutions were prepared by dissolving the peptide in 10 mM potassium phosphate, pH 6.8 containing 15 mM hexamethylphosphoramide, HMPA. Concentrations were calculated using the dissolved weight of the lyophilized peptide.

The DNA sequence for the M1199L mutant peptide, DAD^{M1199L}, was purchased from GenScript and inserted into pTM-7 to express an N-terminal his-tagged TrpL-DAD^{M1199L} fusion protein (35). Fusing the DAD^{M1199L} peptide to a hydrophobic protein (TrpL) directed the gene product into inclusion bodies. The methionine residues in TrpL were mutated to leucine residues (36) to allow for a single cyanogen bromide cleavage site between TrpL and DAD^{M1199L}. The plasmid was transformed into *E. coli* BL21(DE3) to overexpress both labeled and unlabeled peptide as described for the DID constructs.

To purify DAD^{M1199L}, the bacterial cell pellet was resuspended in 50 mM Tris-HCl, pH 7.2, 1% w/v Triton X-100, and 1 mM EDTA, sonicated as described above and centrifuged at 10,000g for 30 min at 4 °C. The pellet was washed with a series of buffers, 50 mM Tris-HCl, pH 7.2, containing 2% w/v Triton X-100 and 1 mM EDTA, followed by 25 mM Tris-HCl, pH 7.2 containing 1 M NaCl and 0.5 mM EDTA, and 50 mM Tris-HCl, pH 7.2. The cells were sonicated and centrifuged between washes. The final pellet was dissolved in denaturing buffer, 50 mM sodium phosphate, pH 8.3 containing 6 M guanidinium chloride, and clarified by centrifugation at 10,000g. The supernatant was incubated with nickel-nitriloacetic acid agarose beads overnight for batch binding. The beads were packed into a column and washed once with denaturing buffer. The column was washed with five volumes of 50 mM sodium phosphate, pH 7 containing 6 M urea, followed by 50 mM sodium phosphate, pH 6.4, and 6 M urea. Elution was carried out using 50 mM sodium phosphate, pH 3.6, and 6 M urea. Fractions containing TrpL-DAD^{M1199L} were pooled, dialyzed into water, and lyophilized. Cleavage was performed by dissolving the TrpL-DAD^{M1199L} fusion protein

in 70% formic acid with a 100 M excess of cyanogen bromide and incubating for 1.5 h at 25 °C. The cleaved products were dried *in vacuo*, loaded onto a C18 column (Agilent ZORBAX 300SB-C18) for HPLC. Cleaved products were resolved using a gradient of 0% to 90% acetonitrile in 0.1% trifluoroacetic acid. DAD^{M1199L} eluted first and was collected, lyophilized, and stored at -20 °C. DAD^{M1199L} stock solutions were prepared by dissolving the peptide in 10 mM potassium phosphate buffer, pH 6.8 containing 15 mM HMPA. The concentration of DAD^{M1199L} stock solutions was determined by integrating HPLC chromatogram traces at 260 nm, the approximate wavelength of maximal absorption of phenylalanine residues relative to reference WT DAD peptide solutions.

T-helix expression and purification

The T-helix protein, KRRETEEKMRRAKLAKEKAEKERLEKQQ, corresponding to residues 1154 to 1181 (UniProt O60610) was purchased from GenScript as a lyophilized powder. Samples were resuspended in 90 mM sodium phosphate, pH 7.2, 75 mM sodium chloride, 0.01% sodium azide, and 10%v/v D₂O at 291 K for structure determination and used immediately for NMR experiments or stored at -20 °C until further use.

Cross-linking reactions

Amine-to-amine cross-linking of unlabeled DID to DAD^{M1199L} was performed using disuccinimidyl glutarate (DSG), bis(sulfosuccinimidyl) suberate (BS3), and bis(succinimidyl) penta(ethylene glycol) (BS(PEG)₅) purchased from Thermo Fisher Scientific. These amine-to-amine cross-linkers had spacer arm lengths of 7.7, 11.4, and 21.7 Å, respectively. Cross-linkers were added to an equimolar mixture of 200 μM DID/DAD^{M1199L}, in cross-linking buffer, 20 mM potassium phosphate buffer, pH 7.5, 100 mM NaCl, to final concentrations of 2.5 mM for DSG and BS3, and 1 mM for BS(PEG)₅. Cross-linking reactions were incubated for 2 h on ice or 30 min at 25 °C. Reactions were quenched by adding 50 mM Tris-HCl, pH 7.5 to a final concentration of 62 mM.

The cross-linked products were resolved by using 8% SDS-PAGE. Bands corresponding to cross-linked DID/DAD^{M1199L} were visualized by staining with Coomassie Blue R-250 (Fig. S3B). The bands of interest were excised from the gels and cut into millimeter-length pieces. The gel pieces were destained in 50% v/v methanol 5% v/v acetic acid, dehydrated by soaking in acetonitrile, and lyophilized. The gel pieces were treated with 10 mM DTT for 15 min at 37 °C to break disulfide bonds, then with 50 mM iodoacetamide for 30 min at 25 °C to alkylate the protein sulfhydryl groups. Excess iodoacetamide was removed by washing the gels with 100 mM NH₄HCO₃, followed by dehydrating in acetonitrile. After three rounds of washing and dehydrating, the gel pieces were completely dried by lyophilization. In-gel tryptic digestion was performed by incubating the lyophilized gel pieces in 100 mM NH₄HCO₃ with 20 ng/μl trypsin/Lys-C (Promega) on ice for 30 min, after which excess liquid was removed and the gel pieces were incubated overnight at 37 °C. Trypsin/Lys-C was chosen for

high fidelity and is expected to only partially cleave the RRKR motif to dicationic or tricationic peptides; tetra-cationic peptides are trypsin digested 100 times faster than dicationic and tricationic regions (56), and thus the full RRKR motif is not expected to remain after tryptic digestion. The gel pieces were washed with 5% v/v formic acid, and peptides were extracted by washing with 5% v/v formic acid and 50% v/v acetonitrile. The extracts were concentrated by lyophilization and stored at -20 °C until use in LC-MS/MS analysis.

Liquid chromatography and MS

MS was performed at the RNA Epitranscriptomics & Proteomics Resource located at the Life Sciences Research Building, University at Albany, supported by the SUNY Research Foundation. LC-MS/MS was performed on an integrated micro LC-Orbitrap Velos system (Thermo Fisher Scientific), comprising a 3-pump Waters Cap LC microscale chromatography system (Waters Corp) with an autosampler, a stream-select module configured for precolumn plus analytical capillary column, and an Orbitrap Velos mass spectrometer fitted with an H-ESI probe operated under xcalibur 2.2 control.

Samples were resuspended in 0.1% v/v formic acid and 3% v/v acetonitrile. The peptides were separated on an ACE C18 capillary column (15 cm × 500 μm internal diameter) packed with ACE C18-300 particles (5 μm resin, Advanced Chromatography Technologies). The C18 column was connected in-line with the mass spectrometer. Peptides were eluted at a flow rate of 20 μl/min with a 40 min gradient of 5 to 80% acetonitrile in 0.1% formic acid into quartz emitters and analyzed by electrospray ionization MS, using a Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) with an emitter voltage of 4.5 kV, a sheath gas flow of 10 and a capillary temperature of 275 °C. The Orbitrap was operated in a data-dependent acquisition mode in which multiple charged ions with abundance >6000 cps were selected for MS/MS fragmentation. Full scan mass spectra, 350 to 2000 *m/z*, with the resolution setting of 30,000 at 200 *m/z*, were detected in the Orbitrap analyzer after accumulation of 10⁶ ions. Peptides with multiple charges were selected automatically and fragmented in the collision cell through high-energy collision dissociation with 34% normalized collision energy at a resolution of 7500. A lock mass of diisooctyl phthalate, *m/z* 391.28428, was used for mass calibration. Data analysis was performed using pLink 2.3.9 (<http://pfind.org/software/pLink/>) software (57). Raw data from the LC-MS/MS experiments were deposited in the jPOSTrepo repository (58) under the accession codes JPST002156 and PXD042130.

Cross-link MS analysis

Raw data from MS were processed using pLink 2.3.9 software (57) using the RAW files and the following settings: flow type, high-energy collision dissociation; enzyme, trypsin, up to five missed sites; precursor mass tolerance, 20 ppm; fragment mass tolerance, 0.02 Da; peptide length, minimum three amino acids and maximum 60 amino acids per chain; peptide mass, minimum 300 and maximum 6000 Da; carbamidomethylation of cysteine, fixed; oxidation of methionine, variable; filter

tolerance, 10 ppm; and false discovery rate <5% at the peptide-spectrum match level. Since only bacterially expressed or chemically synthesized and highly purified proteins were used in this study, the database searched consisted of the DID and DAD sequences and the masses used for cross-linked DSG, BS3, and BS(PEG)₅ were 96.02113 Da, 138.06808 Da and 302.13656 Da, respectively. The results were annotated using pLabel (<http://pfind.org/software/pLabel/index.html>) (59), requiring a mass deviation <0.5 Da.

Fluorescence titrations

Tryptophan fluorescence spectroscopy titration experiments were performed using a Fluorolog-3 Spectrofluorometer (Horiba). Excitation and emission wavelengths were 280 nm and 352 nm, respectively, with a slit width of 5 nm. Temperature was maintained at 20 °C and measurements were performed in triplicate. Protein samples were prepared in binding buffer, 10 mM 3-(N-morpholino)propanesulfonic acid, pH 7.5, 300 mM NaCl, and 0.01% NaN₃, at 200 nM for WT DID and DID^{E326G/E327A} while DAD^{M1199L} concentrations ranged from 1 nM to 15 μM. The dissociation constant, K_D, was calculated using the “one site – total, accounting for ligand depletion” regression model in the GraphPad Prism (<https://www.graphpad.com/>) version 9.0 software (GraphPad) using

$$[\text{DID}-\text{DAD}^{\text{M1199L}}] = [\text{DID}]_{\text{total}} * \frac{[\text{DAD}^{\text{M1199L}}]_{\text{free}}}{\left(K_D + [\text{DAD}^{\text{M1199L}}]_{\text{free}} \right) + \alpha [\text{DAD}^{\text{M1199L}}]_{\text{free}}}$$

where [DID-DAD^{M1199L}] is the amount of bound complex, K_D is the dissociation constant, and α is the nonspecific-binding coefficient (60). Through algebraic manipulations, the following quadratic is defined

$$\begin{aligned} (F - F_0) / F_0 = \beta & \left[(K_D + [\text{DAD}^{\text{M1199L}}]) (1 + \alpha) + [\text{DID}]_{\text{total}} \right. \\ & + \alpha [\text{DAD}^{\text{M1199L}}] + \left. \left[((K_D + [\text{DAD}^{\text{M1199L}}]) (1 + \alpha) + [\text{DID}]_{\text{total}} \right. \right. \\ & \left. \left. + \alpha [\text{DAD}^{\text{M1199L}}])^2 - 4(1 + \alpha) ([\text{DAD}^{\text{M1199L}}] \right. \right. \\ & \left. \left. (\alpha (K_D + [\text{DAD}^{\text{M1199L}}]) + [\text{DID}]_{\text{total}})) \right]^{1/2} \right] / [2(1 + \alpha)] \end{aligned}$$

where F is the fluorescence at a specific DAD^{M1199L} concentration, F₀ is the fluorescence intensity of the blank, β is a scaling factor in the units of μM, [DAD^{M1199L}] is the total ligand concentration and α = 0.00001.

Stopped flow kinetics

Stopped flow fluorescence measurements were performed at 20 °C using a Stopped-Flow Module SFM-400 (Bio-Logic)

DIAPH1 and the productive encounter complex

equipped with an MPS-60 motor power supply, an ALX-250 arc lamp and a PMS-250 photo multiplier system. The tryptophan residue of DID was excited at 285 nm, and the fluorescence emission was measured at 305 nm with 700 V applied to the photo multiplier tube. Two hundred nanomolar (200 nM) DID, in binding buffer, was titrated from 1.5 μM to 20 μM with DAD^{M1199L}. The measured voltage was plotted as a function of time and fit to a single exponential using a pseudo first-order kinetic model to estimate k_{obs} , where $k_{\text{obs}} = k_{\text{on}}[\text{DAD}^{\text{M1199L}}] + k_{\text{off}}$. Linear regression of k_{obs} versus $[\text{DAD}^{\text{M1199L}}]$ using Prism v9 software (GraphPad) yielded estimates for k_{on} (slope) and k_{off} (y-intercept), where k_{on} is the rate of complex formation and k_{off} is the rate of complex dissociation.

NMR spectroscopy

All experiments were performed in 20 mM potassium phosphate, pH 7.5, 100 mM sodium chloride, 1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride, 0.5 mM EDTA, 1 mM DTT, 15 mM HMPA, and 10% v/v D₂O unless stated otherwise. NMR spectra were acquired using a 700 MHz Bruker Avance II or a 600 MHz Bruker Ascend spectrometer equipped with ultrasensitive cryoprobes. Experiments were performed with Watergate (24) water suppression, processed using TopSpin (<https://www.bruker.com/en/products-and-solutions/mr/nmr-software/topspin.html>) version 3.2 or 3.6, and assigned using Computer Aided Resonance Assignment (CARA, <http://cara.nmr.ch>) (25).

Three dimensional HNCA, HNCACB, CBCACONH and HNCO as well as 2D ¹H-¹⁵N HSQC (38), spectra were recorded at 305 K to assign backbone residues of 250 μM [^U-¹⁵N, ¹³C]-DID in the presence of a molar excess of WT DAD. To facilitate assignment of DID residues, 2D ¹H-¹⁵N HSQC and 3D ¹⁵N NOESY spectra of DID bound to WT DAD with DID containing selectively ¹⁵N-labeled amino acids were recorded at 305 K or 298 K to facilitate DID resonance assignment. Selective labeling was accomplished for Ala, Asn, Met, Tyr, Cys, Arg, Gly, His, Ile, Leu, Lys, Phe, Thr, Trp, and Val.

Three dimensional HNCA and 2D ¹H-¹⁵N HSQC spectra were recorded with 300 μM [^U-¹⁵N, ¹³C]-DID in the presence of 350 μM DAD^{M1199L} to facilitate transfer of H^N, N, and CA assignments from DID-DAD spectra to DID-DAD^{M1199L} spectra. Side-chain proton assignments of [^U-¹⁵N, ¹³C]-DID bound to DAD^{M1199L} were obtained by recording H(CCCO)NH and 3D ¹⁵N-NOESY spectra (38). Backbone and side-chain assignments of free [^U-¹⁵N, ¹³C]-DAD^{M1199L} at 298 K were determined by recording (H)CCH-TOCSY, 3D CBCACONH, 2D ¹H-¹³C HSQC, 2D ¹H-¹⁵N HSQC, 3D HNCA, and 3D HNCOCA spectra (38). Backbone and side-chain assignments of [^U-¹⁵N, ¹³C]-DAD^{M1199L} bound to DID were determined by recording 3D HNCA, and 3D HNCOCA, (H)CCH-TOCSY and 3D HBHACONH at 305 K (38).

To determine temperature dependence of the transition between DID conformers A and B, 2D ¹H-¹⁵N HSQC spectra were recorded with [^U-¹⁵N]-DID bound to DAD^{M1199L} at 280,

283, 285, 288, 290, 292, 295, 298, 300, 303, 305, 308, and 310 K. The peak amplitudes of DID C-terminus residues G369, S373, Y374, and G378 were quantified. Equilibrium constants (K_{eq}) were calculated for the transition between conformer A and B as the ratio of the cross-peak amplitudes of conformer A over conformer B. A linear fit of the van't Hoff plot, $\ln(K_{\text{eq}})$ versus $1/T$ (K⁻¹) was performed using GraphPad Prism 9.0 software. The linear regression analysis yielded thermodynamic parameters ΔG , ΔH , and ΔS for the conformer transition.

3D ¹⁵N-NOESY spectra were collected for [^U-¹⁵N]-DID in the presence of excess unlabeled DAD^{M1199L} and for [^U-¹⁵N]-DAD^{M1199L} in the presence of unlabeled DID. To facilitate the identification of intermolecular nOes, 3D ¹⁵N-NOESY spectra were collected for [^U-¹⁵N]-DID in the presence of [^U-¹³C]-DAD^{M1199L} in the presence and absence of a decoupling pulse on the ¹³C channel during ¹H detection in the Bruker pulse program noesyhsqc3gp193d.

Steady state ¹H-¹⁵N nOe spectra (61) were collected at 305 K using 450 μM [^U-¹⁵N] DID with 650 μM unlabeled DAD^{M1199L} and 100 μM [^U-¹⁵N]-DAD^{M1199L} with 200 μM unlabeled DID. Measurements of ¹H-¹⁵N nOes for free [^U-¹⁵N]-DAD^{M1199L} were performed at 298 K to reduce exchange broadening. Interleaved spectra were recorded with and without proton saturation through a high-power 120° pulse applied every 5 ms within the 5 s recycle delay. Quantification of ¹H-¹⁵N nOes was performed using the equation $\{^1\text{H}\}\text{-}^{15}\text{N nOe} = I_{\text{saturated}}/I_{\text{unsat}}$, where $I_{\text{saturated}}$ is the peak amplitude from spectra collected with ¹H saturation, and I_{unsat} corresponds to the condition without saturation. Error estimates of the ¹H-¹⁵N nOe, δ_{nOe} , were determined from the noise levels in saturated, $\delta_{\text{saturated}}$, and unsaturated spectra, δ_{unsat} , using the equation

$$\delta_{\text{nOe}} / \text{nOe} = [(\delta_{\text{saturated}}/I_{\text{saturated}})^2 + (\delta_{\text{unsat}}/I_{\text{unsat}})^2]^{1/2}$$

NMR structure calculations

CYANA²⁶ version 3.98.5 was used to calculate the 3D structure of DID in complex with DAD^{M1199L}. The input consisted of a chemical shift list obtained from the resonance assignment, a sequence file of a single polypeptide chain in which the DID sequence was connected to DAD^{M1199L} via linker residues (LL5 in the standard CYANA library), an unassigned 3D ¹⁵N-NOESY peak list for DID, an unassigned 3D ¹⁵N-NOESY peak list for DAD^{M1199L}, upper distance limits from cross-linking MS data (Table S3 footnote b), upper distance limits for manually assigned intramolecular and intermolecular nOes set at 5 Å, and distance restraints for helical regions identified in crystal structure PDB 2F31. Hydrogen bond N-O distances in acceptor-donor pairs were relaxed by 0.5 Å and N-C distances in helical regions were relaxed by 0.5 Å.

NOESY peaks were assigned automatically and converted into distance constraints using the standard CYANA macro with seven cycles of nOe assignment and simulated annealing in torsion angle space. The I-PINE NMR server (<http://pine>).

nmrfam.wisc.edu) classified Pro 272 of DID as *cis* and all other DID prolines as *trans* based on the average C β chemical shifts (62), therefore Pro 272 was designated as cPro in the sequence file. The CISPROCHECK routine of CYANA 3.98.5 classified the isomerization state of Pro 1218 in DAD^{M1199L} as *trans* based on the average value and SD for the difference between the chemical shifts of C β and C γ .²⁹ Backbone Φ and Ψ dihedral angle constraints were determined using the TALOS-N webserver, (<http://spin.niddk.nih.gov/bax/software/TALOS-N>) (40). The experimental restraints used for the structure calculation are summarized in Tables S2 and S3.

CYANA-generated structures were subjected to minimization in explicit water using YASARA Structure version 20.12.24 (63). Distance restraints from manual and automatic assignment of nOes and cross-linking MS as well as the torsion angle list used in CYANA simulated annealing were used as input in the refinement procedure. In YASARA Structure version 20.12.24, refinement was performed using the Amber14 force field and the *nmr_refinesolvent* macro. Coordinates of the refined ensemble of 20 DID-DAD^{M1199L} structures were deposited in the Protein Data Bank with entry code 8FG1 (<https://deposit-1.wwpdb.org/>).

Chemical shift index analysis

DID, DAD^{M1199L}, and T-helix secondary structures were predicted using chemical shift index, CSI, analyses on the CSI 3.0 web server <http://csi3.wishartlab.com>. The consensus secondary structures were defined in terms of probabilities for a particular segment in DID/DAD^{M1199L} or T-helix to fold into α -helix, random coil, or β -sheet.

Iso-surface map visualization

Iso-surface maps of the DID-DAD^{M1199L} and DID^{E326G/E327A}-DAD^{M1199L} complexes were constructed using an optimized DID-DAD^{M1199L} structure from our refinements. Using the Adaptive Poisson-Boltzmann Solver (64) (<https://server.poissonboltzmann.org/>), the electronic surface properties of the proteins were calculated using the AMBER force field. Results were visualized using the Visual Molecular Dynamics (65) (<https://www.ks.uiuc.edu/Research/vmd/>) software with blue surfaces set to +10 eV and red surfaces set to -10 eV.

Construction of plasmids for colocalization experiments

Preparation of plasmid constructs expressing human DIAPH1 and DIAPH1 ^{Δ DAD} with C-terminal monomeric enhanced YFP fluorescent tags, DIAPH1-YFP and DIAPH1 ^{Δ DAD}-YFP, was described previously (10, 66). Constitutively active DIAPH1 ^{Δ DAD} was obtained by deleting the C-terminal DAD, which corresponds to amino acids 1194 to 1272 of DIAPH1. YFP fluorescent tags were excised from the *AgeI* and *NotI* sites of pCAG-YFP (Addgene) and GFP was inserted resulting in plasmids pDIAPH1-GFP and pDIAPH1 ^{Δ DAD}-GFP. To obtain DIAPH1 constructs with N-terminal YFP tags, YFP was excised at the *AgeI* and *NotI* sites,

and the gap was closed using a short linker duplex, 5'-CCGGTGTAAATAGTGC-3', 5'-GGCCGCGCACTATTACA-3', containing two stop codons. YFP was amplified with two oligonucleotide primers 5'-TTAAGCGCTATGGTGAGCAAGGGCGAGGAG-3' and 5'-TTACTCGAGACCGCCGCTACCGCCCTTGTACAGCTCGTCCATGCCGAG-3', digested, and cloned into the *AfeI* and *XhoI* sites in upstream of the DIAPH1 sequence resulting in plasmid pYFP-DIAPH1. Self-complementary oligonucleotides 5'-GCTCTCATCACCA GCGGGCGCCCTTGACTTCCGAGTTCACA-3' and 5'-TGTGAAGCTCGGAAGTCAAGGGCGCCCGCTGGTGTGATGAGAGC-3' were used with the QuickChange Lightning site-directed Mutagenesis Kit (Agilent Technologies) to introduce mutations E326G/E327A into the DIAPH1 sequence resulting in plasmids pYFP-DIAPH1^{E326G/E327A} and pDIAPH1^{E326G/E327A}-GFP.

HEK 293T cell culture, transfection, confocal microscopy

HEK 293T cells were grown in Dulbecco's modified Eagle's medium supplemented with 4 mM L-glutamine, 5.5 mM D-glucose, 1 mM sodium pyruvate, 10% fetal bovine serum (FBS), (HyClone), 100 units/ml penicillin, and 100 mg/ml of streptomycin. Fifteen millimeter (15 mm) diameter round glass coverslips (Ted Pella) were coated with Collagen I and seeded with 5×10^4 cells each in the wells of 24-well plate. After incubating overnight at 37 °C, the Dulbecco's modified Eagle's medium was exchanged with Opti-minimal essential medium/5% FBS and the cells were transfected with plasmids using FuGENE 6 reagent according to the manufacturer's recommendations. After 48 h of growth, cells were fixed by replacing the medium with 4% formaldehyde in PBS for 20 min at room temperature and washed three times with PBS. Cells were incubated with Alexa Fluor 568 Phalloidin (Thermo Fisher Scientific) solution in PBS, prepared in line with the manufacturer's guidelines, for 45 min at 37 °C, washed three times with PBS, and mounted on slides using Fluoromount G (Electron Microscopy Sciences).

Images were acquired by using a Zeiss LSM 710 confocal imaging system (Carl Zeiss Microscopy) equipped with an X63, 1.4 NA oil immersion objective. GFP was excited with a 488 nm argon laser line at 1% power and Alexa Fluor 568 with a 561 nm line at 0.8% power. Image stacks were recorded at 290 nm steps along the z-axis in two-channel mode, using a 488/561 nm beam splitter with a 492 to 550 nm emission window for the GFP channel and a 592 to 676 nm window for the Alexa Fluor 568 channel. A total of eight Z-stacks were acquired for DIAPH1 ^{Δ DAD}-GFP/F-Actin cells, and nine each for DIAPH1^{E326G/E327A}-GFP/F-Actin and DIAPH1-GFP/F-Actin cells.

3D images were processed using FluoRender 2.27 software (67) (<https://www.sci.utah.edu/software/fluorender.html>). ImageJ plugin JACoP 2.0 (68) (<https://imagej.nih.gov/ij/plugins/track/jacop2.html>) was used to assess the degree of colocalization of F-Actin with DIAPH1 or its mutants. Z-stacks were processed at default settings in 3D mode, and Manders coefficients M1 and M2, which measures the amount of colocalization for each image were recorded. Colocalization data was processed with Prism

DIAPH1 and the productive encounter complex

9.0.2 (<https://www.graphpad.com/>) software (GraphPad). Column heights reported on the graph represent mean values, error bars represent the SD. The statistical significance of multiple comparisons was estimated by using the one-way ANOVA subroutine and Fisher's least significant difference test with a single pooled variance displayed in GraphPad style, where the absence of bars represents $p > 0.05$, * is $p \leq 0.05$, ** is $p \leq 0.01$, and *** is $p \leq 0.001$.

Human primary aortic vascular SMC migration

Human primary aortic vascular SMCs (ATCC PCS-100-012) were purchased from American Type Tissue Culture Collection. Cells were propagated in Vascular Cell Basal Medium (ATCC PCS-100-030) supplemented with the Vascular Smooth Muscle Cell Growth Kit (ATCC PCS-100-042) containing recombinant human, rh-FGF-b, rh-insulin, rh-EGF, ascorbic acid, L-glutamine, FBS, and 0.1% penicillin/streptomycin in a humidified atmosphere of 5% CO₂ at 37 °C.

Lipofectamine LTX reagent (Thermo Fisher Scientific) was used to perform transfection according to manufacturer's protocol. The protocol was optimized for 2.5 µg/ml of plasmid DNA in reduced serum Opti-minimal essential medium for efficient transfection. Human primary aortic SMCs were assessed for migration in response to 10 µg/ml of RAGE ligand, CML-HSA (8) with a wounding assay as previously described (69). Briefly, cells were grown to confluence in 24-well plates and serum-starved overnight. The following morning, the monolayer was wounded using a p10 pipette tip and fresh medium containing RAGE ligand was added for 4 h. Cells were maintained at 37 °C and 5% CO₂. Images were taken at 4 h intervals. Each image was measured, and an ingrowth area of effective migrating cells was calculated. The data was subjected to a Grubbs test with $\alpha = 0.1$, and the resulting data was processed with Prism v 9.0.2 software (GraphPad). Column heights reported on the graph represent mean values, error bars represent the SD. The statistical significance of multiple comparisons was estimated by using the one-way ANOVA subroutine.

Data availability

- Expression plasmids generated in this study are available upon request from the corresponding author with a completed materials transfer agreement.
- Chemical shift assignments were deposited into the Biological Magnetic Resonance Bank with accession code: 31064.
- Distance restraints and atomic coordinates for reported structures are deposited into the Protein Data Bank with accession code: 8FG1.
- MS data was deposited to <https://repository.jpostdb.org/> under acquisition codes JPST002156 and PXD042130.

Supporting information—This article contains supporting information (57, 60, 70, 71).

Acknowledgments—The project described was supported by Award Numbers 1P01HL146367-01, 5R01DK122456 and R01GM085006 from the U.S. National Institute of Health. Mass spectrometry analyses were performed by the RNA Epitranscriptomics & Proteomics Resource located at the Life Science Research Building, University at Albany, supported by the SUNY Research Foundation and NIH Award 5R01GM143749.

Author contributions—G. G. T., L. M. S. R., A. P., S. R., M. B. M., and G. Y. investigation; G. G. T., L. M. S. R., A. P., S. R., M. B. M., and G. Y. formal analysis. R. R., A. M. S., and A. S. conceptualization; G.G.T., L.M.S. R., D. S. B., G.Y., R. R., A. M. S., and A. S. writing-original draft; G. G. T., L. M. S. R., A. P., S. R., M. B. M., G. Y., D. S. B., R. R., A. M. S., and A. S. writing-review and editing.

Funding and additional information—The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—The authors declare that they have no conflicts of interest with the content of this article.

Abbreviations—The abbreviations used are: ARR, armadillo repeat region; BS3, bis(sulfosuccinimidyl) suberate; BS(PEG)₅, bis(succinimidyl) penta(ethylene glycol); CML-HAS, carboxymethyllysine human serum albumin; CSI, chemical shift index; DAD, diaphanous autoinhibitory domain; DIAPH1, Diaphanous 1; DID, diaphanous inhibitory domain; DRF, diaphanous-related formin; DSG, disuccinimidyl glutarate; FBS, fetal bovine serum; FH, formin homology; HMPA, hexamethylphosphoramide; HSQC, heteronuclear single-quantum coherence; IH, interdomain helix; MS, mass spectrometry; MS/MS, tandem mass spectrometry; RAGE, receptor for advanced glycation end products; SMC, smooth muscle cell; YFP, yellow fluorescent protein.

References

1. Pollard, T. D., and Borisy, G. G. (2003) Cellular motility driven by assembly and disassembly of actin filaments. *Cell* **112**, 453–465
2. Pelham, R. J., and Chang, F. (2002) Actin dynamics in the contractile ring during cytokinesis in fission yeast. *Nature* **419**, 82–86
3. Copeland, J. W., and Treisman, R. (2002) The diaphanous-related formin mDia1 controls serum response factor activity through its effects on actin polymerization. *Mol. Biol. Cell* **13**, 4088–4099
4. Higashida, C., Miyoshi, T., Fujita, A., Ocegüera-Yanez, F., Monypenny, J., Andou, Y., *et al.* (2004) Actin polymerization-driven molecular movement of mDia1 in living cells. *Science* **303**, 2007–2010
5. Tominaga, T., Sahai, E., Chardin, P., McCormick, F., Courtneidge, S. A., and Alberts, A. S. (2000) Diaphanous-related formins bridge Rho GTPase and Src tyrosine kinase signaling. *Mol. Cell* **5**, 13–25
6. Kato, T., Watanabe, N., Morishima, Y., Fujita, A., Ishizaki, T., and Narumiya, S. (2001) Localization of a mammalian homolog of diaphanous, mDia1, to the mitotic spindle in HeLa cells. *J. Cell Sci.* **114**, 775–784
7. Toure, F., Fritz, G., Li, Q., Rai, V., Daffu, G., Zou, Y. S., *et al.* (2012) Formin mDia1 mediates vascular remodeling via integration of oxidative and signal transduction pathways. *Circ. Res.* **110**, 1279–1293
8. Hudson, B. I., Kalea, A. Z., Del Mar Arriero, M., Harja, E., Boulanger, E., D'Agati, V., *et al.* (2008) Interaction of the RAGE cytoplasmic domain with diaphanous-1 is required for ligand-stimulated cellular migration through activation of Rac1 and Cdc42. *J. Biol. Chem.* **283**, 34457–34468
9. Rai, V., Maldonado, A. Y., Burz, D. S., Reverdatto, S., Yan, S. F., Schmidt, A. M., *et al.* (2012) Signal transduction in receptor for advanced glycation end products (RAGE): solution structure of C-terminal rage (ctRAGE) and its binding to mDia1. *J. Biol. Chem.* **287**, 5133–5144

10. Xue, J., Manigrasso, M., Scalabrin, M., Rai, V., Reverdatto, S., Burz, D. S., *et al.* (2016) Change in the molecular dimension of a RAGE-ligand complex triggers RAGE signaling. *Structure* **24**, 1509–1522
11. Lammers, M., Rose, R., Scrima, A., and Wittinghofer, A. (2005) The regulation of mDia1 by autoinhibition and its release by Rho*GTP. *EMBO J.* **24**, 4176–4187
12. Li, F., and Higgs, H. N. (2003) The mouse Formin mDia1 is a potent actin nucleation factor regulated by autoinhibition. *Curr. Biol.* **13**, 1335–1340
13. Higgs, H. N. (2005) Formin proteins: a domain-based approach. *Trends Biochem. Sci.* **30**, 342–353
14. Schonichen, A., and Geyer, M. (2010) Fifteen formins for an actin filament: a molecular view on the regulation of human formins. *Biochim. Biophys. Acta* **1803**, 152–163
15. Otomo, T., Tomchick, D. R., Otomo, C., Machius, M., and Rosen, M. K. (2010) Crystal structure of the Formin mDia1 in autoinhibited conformation. *PLoS One* **5**, e12896
16. Nezami, A. G., Poy, F., and Eck, M. J. (2006) Structure of the autoinhibitory switch in formin mDia1. *Structure* **14**, 257–263
17. Nezami, A., Poy, F., Toms, A., Zheng, W., and Eck, M. J. (2010) Crystal structure of a complex between amino and carboxy terminal fragments of mDia1: insights into autoinhibition of diaphanous-related formins. *PLoS One* **5**, e12992
18. Yamada, Y., and Liu, D. X. (2009) Proteolytic activation of the spike protein at a novel RRRR/S motif is implicated in Furin-dependent entry, Syncytium formation, and Infectivity of coronavirus infectious Bronchitis Virus in cultured cells. *J. Virol.* **83**, 8744–8758
19. Siegfried, G., Khatib, A. M., Benjannet, S., Chretien, M., and Seidah, N. G. (2003) The proteolytic processing of pro-platelet-derived growth factor-A at RRRR86 by members of the proprotein convertase family is functionally correlated to platelet-derived growth factor-A-induced functions and tumorigenicity. *Cancer Res.* **63**, 1458–1463
20. Holdermann, I., Meyer, N. H., Round, A., Wild, K., Sattler, M., and Sinning, I. (2012) Chromodomains read the arginine code of post-translational targeting. *Nat. Struct. Mol. Biol.* **19**, 260–263
21. Fujiwara, Y., Fujiwara, K., Goda, N., Iwaya, N., Tenno, T., Shirakawa, M., *et al.* (2011) Structure and function of the N-terminal Nucleolin binding domain of nuclear valosin-containing protein-like 2 (NVL2) Harboring a Nucleolar localization signal. *J. Biol. Chem.* **286**, 21732–21741
22. Wallar, B. J., Stropich, B. N., Schoenherr, J. A., Holman, H. A., Kitchen, S. M., and Alberts, A. S. (2006) The basic region of the diaphanous-autoregulatory domain (DAD) is required for autoregulatory interactions with the diaphanous-related formin inhibitory domain. *J. Biol. Chem.* **281**, 4300–4307
23. Lammers, M., Meyer, S., Kuhlmann, D., and Wittinghofer, A. (2008) Specificity of interactions between mDia isoforms and Rho proteins. *J. Biol. Chem.* **283**, 35236–35246
24. Ueyama, T., Ninoyu, Y., Nishio, S. Y., Miyoshi, T., Torii, H., Nishimura, K., *et al.* (2016) Constitutive activation of DIA1 (DIAPH1) via C-terminal truncation causes human sensorineural hearing loss. *EMBO Mol. Med.* **8**, 1310–1324
25. Rose, R., Weyand, M., Lammers, M., Ishizaki, T., Ahmadian, M. R., and Wittinghofer, A. (2005) Structural and mechanistic insights into the interaction between Rho and mammalian Dia. *Nature* **435**, 513–518
26. O'Shea, K. M., Ananthakrishnan, R., Li, Q., Quadri, N., Thiagarajan, D., Sreejit, G., *et al.* (2017) The formin, DIAPH1, is a key modulator of myocardial ischemia/reperfusion injury. *EBioMedicine* **26**, 165–174
27. Manigrasso, M. B., Friedman, R. A., Ramasamy, R., D'Agati, V., and Schmidt, A. M. (2018) Deletion of the formin Diaph1 protects from structural and functional abnormalities in the murine diabetic kidney. *Am. J. Physiol. Renal Physiol.* **315**, F1601–F1612
28. Eisenmann, K. M., West, R. A., Hildebrand, D., Kitchen, S. M., Peng, J., Sigler, R., *et al.* (2007) T cell responses in mammalian diaphanous-related formin mDia1 knock-out mice. *J. Biol. Chem.* **282**, 25152–25158
29. Ercan-Sencicek, A. G., Jambi, S., Franjic, D., Nishimura, S., Li, M., El-Fishawy, P., *et al.* (2015) Homozygous loss of DIAPH1 is a novel cause of microcephaly in humans. *Eur. J. Hum. Genet.* **23**, 165–172
30. Kaustio, M., Nayebzadeh, N., Hinttala, R., Tapiainen, T., Astrom, P., Mamia, K., *et al.* (2021) Loss of DIAPH1 causes SCBMS, combined immunodeficiency, and mitochondrial dysfunction. *J. Allergy Clin. Immunol.* **148**, 599–611
31. Lakha, R., Montero, A. M., Jabeen, T., Costeas, C. C., Ma, J., and Vizcarra, C. L. (2021) Variable autoinhibition among deafness-associated variants of diaphanous 1 (DIAPH1). *Biochemistry* **60**, 2320–2329
32. Stritt, S., Nurden, P., Turro, E., Greene, D., Jansen, S. B., Westbury, S. K., *et al.* (2016) A gain-of-function variant in DIAPH1 causes dominant macrothrombocytopenia and hearing loss. *Blood* **127**, 2903–2914
33. Tang, C., Iwahara, J., and Clore, G. M. (2006) Visualization of transient encounter complexes in protein-protein association. *Nature* **444**, 383–386
34. Andralojc, W., Hiruma, Y., Liu, W. M., Ravera, E., Nojiri, M., Parigi, G., *et al.* (2017) Identification of productive and futile encounters in an electron transfer protein complex. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1840–E1847
35. Staley, J. P., and Kim, P. S. (1994) Formation of a native-like subdomain in a partially folded intermediate of bovine pancreatic trypsin inhibitor. *Protein Sci.* **3**, 1822–1832
36. Ramirez, L. S., Pande, J., and Shekhtman, A. (2019) Helical structure of recombinant melittin. *J. Phys. Chem. B* **123**, 356–368
37. French, S., and Robson, B. (1983) What is a conservative substitution. *J. Mol. Evol.* **19**, 171–175
38. Cavanagh, J., Skelton, N. J., Fairbrother, W. J., Rance, M., and Palmer, A. G. (2007) *Protein NMR Spectroscopy*, Academic Press, New York, NY
39. Fushman, D., Cahill, S., and Cowburn, D. (1997) The main-chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: analysis of N-15 relaxation with monomer/dimer equilibration. *J. Mol. Biol.* **266**, 173–194
40. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289–302
41. Guntert, P., and Buchner, L. (2015) Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol. NMR* **62**, 453–471
42. Krieger, E., and Vriend, G. (2014) YASARA View-molecular graphics for all devices-from smartphones to workstations. *Bioinformatics* **30**, 2981–2982
43. Otomo, T., Otomo, C., Tomchick, D. R., Machius, M., and Rosen, M. K. (2005) Structural basis of Rho GTPase-mediated activation of the formin mDia1. *Mol. Cell* **18**, 273–281
44. Brenig, J., de Boor, S., Knyphausen, P., Kuhlmann, N., Wroblowski, S., Baldus, L., *et al.* (2015) Structural and biochemical basis for the inhibitory effect of Liprin-alpha 3 on mouse diaphanous 1 (mDia1) function. *J. Biol. Chem.* **290**, 14314–14327
45. Xue, Y., Yuwen, T., Zhu, F., and Skrynnikov, N. R. (2014) Role of electrostatic interactions in binding of peptides and intrinsically disordered proteins to their folded targets. 1. NMR and MD characterization of the complex between the c-Crk N-SH3 domain and the peptide Sos. *Biochemistry* **53**, 6473–6495
46. Ahmad, M., Gu, W., and Helms, V. (2008) Mechanism of fast peptide recognition by SH3 domains. *Angew. Chem. Int. Ed Engl.* **47**, 7626–7630
47. Dunn, K. W., Kamocka, M. M., and McDonald, J. H. (2011) A practical guide to evaluating colocalization in biological microscopy. *Am. J. Physiol. Cell Physiol.* **300**, C723–C742
48. Maiti, S., Michelot, A., Gould, C., Blanchoin, L., Sokolova, O., and Goode, B. L. (2012) Structure and activity of full-length formin mDia1. *Cytoskeleton (Hoboken)* **69**, 393–405
49. Schaks, M., Giannone, G., and Rottner, K. (2019) Actin dynamics in cell migration. *Essays Biochem.* **63**, 483–495
50. Machacek, M., Hodgson, L., Welch, C., Elliott, H., Pertz, O., Nalbant, P., *et al.* (2009) Coordination of Rho GTPase activities during cell protrusion. *Nature* **461**, 99–103
51. Bustos, R. I., Forget, M. A., Settleman, J. E., and Hansen, S. H. (2008) Coordination of Rho and rac GTPase function via p190B RhoGAP. *Curr. Biol.* **18**, 1606–1611
52. Li, Z. L., and Buck, M. (2019) Modified potential functions result in enhanced predictions of a protein complex by all-atom molecular dynamics simulations, confirming a stepwise association process for native protein-protein interactions. *J. Chem. Theory Comput.* **15**, 4318–4331

DIAPH1 and the productive encounter complex

53. Sugase, K., Dyson, H. J., and Wright, P. E. (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025
54. Gould, C. J., Maiti, S., Michelot, A., Graziano, B. R., Blanchoin, L., and Goode, B. L. (2011) The formin DAD domain plays dual roles in auto-inhibition and actin nucleation. *Curr. Biol.* **21**, 384–390
55. McIntosh, L. P., and Dahlquist, F. W. (1990) Biosynthetic incorporation of ¹⁵N and ¹³C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. *Q. Rev. Biophys.* **23**, 1–38
56. Waley, S. G., and Watson, J. (1953) The action of trypsin on polylysine. *Biochem. J.* **55**, 328–337
57. Chen, Z. L., Meng, J. M., Cao, Y., Yin, J. L., Fang, R. Q., Fan, S. B., *et al.* (2019) A high-speed search engine pLink 2 with systematic evaluation for proteome-scale identification of cross-linked peptides. *Nat. Commun.* **10**, 3404
58. Okuda, S., Watanabe, Y., Moriya, Y., Kawano, S., Yamamoto, T., Matsumoto, M., *et al.* (2017) jPOSTrepo: an international standard data repository for proteomes. *Nucleic Acids Res.* **45**, D1107–D1111
59. Wang, L. H., Li, D. Q., Fu, Y., Wang, H. P., Zhang, J. F., Yuan, Z. F., *et al.* (2007) pFind 2.0: a software package for peptide and protein identification via tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **21**, 2985–2991
60. Swillens, S. (1995) Interpretation of binding curves obtained with high receptor concentrations: practical aid for computer analysis. *Mol. Pharmacol.* **47**, 1197–1203
61. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., *et al.* (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ¹⁵N NMR relaxation. *Biochemistry* **33**, 5984–6003
62. Shen, Y., and Bax, A. (2010) Prediction of Xaa-Pro peptide bond conformation from sequence and chemical shifts. *J. Biomol. NMR* **46**, 199–204
63. Krieger, E., and Vriend, G. (2002) Models@home: distributed computing in bioinformatics using a screensaver based approach. *Bioinformatics* **18**, 315–318
64. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10037–10041
65. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38
66. Xie, J., Reverdatto, S., Frolov, A., Hoffmann, R., Burz, D. S., and Shekhtman, A. (2008) Structural basis for pattern recognition by the receptor for advanced glycation end products (RAGE). *J. Biol. Chem.* **283**, 27255–27269
67. Wan, Y., Otsuna, H., Chien, C. B., and Hansen, C. (2012) FluoRender: an application of 2D image space methods for 3D and 4D confocal microscopy data visualization in neurobiology research. *IEEE Pac. Vis. Symp.*, 201–208. <https://doi.org/10.1109/pacificvis.2012.6183592>
68. Bolte, S., and Cordelieres, F. P. (2006) A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213–232
69. Manigrasso, M. B., Pan, J., Rai, V., Zhang, J., Reverdatto, S., Quadri, N., *et al.* (2016) Small molecule inhibition of ligand-stimulated RAGE-DIAPH1 signal transduction. *Sci. Rep.* **6**, 22450
70. Wishart, D. S., and Sykes, B. D. (1994) The ¹³C chemical-shift index: a simple method for the identification of protein secondary structure using ¹³C chemical-shift data. *J. Biomol. NMR* **4**, 171–180
71. Beveridge, R., Stadlmann, J., Penninger, J. M., and Mechtler, K. (2020) A synthetic peptide library for benchmarking crosslinking-mass spectrometry search engines for proteins and protein complexes. *Nat. Commun.* **11**, 742
72. Shimada, A., Nyitrai, M., Vetter, I. R., Kuhlmann, D., Bugyi, B., Narumiya, S., *et al.* (2004) The core FH2 domain of diaphanous-related formins is an elongated actin binding protein that inhibits polymerization. *Mol. Cell* **13**, 511–522