**Structure *18***

# Supplemental Information

# Formation of Salt Bridges Mediates Internal

# Dimerization of Myosin VI Medial Tail Domain

**HyeongJun Kim, Jen Hsin, Yanxin Liu, Paul R. Selvin, and Klaus Schulten**

## SUPPLEMENTAL EXPERIMENTAL PROCEDURES

### Coarse-grained molecular dynamics

Since there is no structural information for the myosin VI medial tail (MT) domain in the dimerized conformation, application of molecular dynamics (MD) methodologies first aimed at establishing a viable model of the MT domain dimer. The following series of protocols was adopted to establish dimer models for two constructs, one involving residues 907 to 980 and one involving residues 907 to 940.

Self-association of the myosin VI MT domains was brought about in an MD simulation based on a simplified representation, residue-based coarse graining (RBCG) (Shih *et al*. 2006. cited in the main text). Usage of a simplified representation, rather than a full all-atom representation, is dictated by the micro-second timescale typically required for spontaneous assembly of biomolecules (Shih *et al*. 2007a and 2007b. cited in the main text), which is near the current computational limit [1]. The all-atom structure of the MT domain segment was first transformed into the RBCG representation, with each amino acid described by two beads (the only exception being glycine, described by one bead only). Two MT segments were then placed in a water box, which, along with neutralizing ions at 50 mM, was also represented in the RBCG form (Shih *et al*. 2006. cited in the main text). The initial separation of the MT segments was ~30 Å. The RBCG systems were simulated using the MD package NAMD (Phillips *et al*. 2005. cited in the main text) with parameters adopted from Shih *et al*. (2006. cited in the main text).

### Recovering all-atom resolution from the coarse-grained representation

The RBCG simulations of both constructs of MT domain (Simulations CG-2MT-907-to-980 and CG-2MT-907-to-940 in Table SS1) displayed the tendency of MT domains to self-associate. To discern the molecular mechanism that stabilizes the helix-helix association, all-atom resolution was recovered from the dimerized RBCG systems.

For this purpose, the RBCG beads representing the proteins were first mapped back to all-atom amino acids, with water and counter ions (again at 50 mM concentration) added, following the reverse coarse graining procedure described in [2]. However, the secondary structure of the MT helices was not fully regained and a subsequent all-atom MD quickly led to disrupted protein chains.

To repair the integrity of the secondary structure of the MT helices, we employed the molecular dynamics flexible fitting (MDFF) method (Trabuco *et al*. 2008 and 2009; Wells *et al*. 2007. all cited in the main text). A 3Å-resolution density map was generated for the annealed MT segments using our program VMD [3]. Subsequently, two equilibrated, all-atom, MT segments generated in Simulation Eq-1MT-907-to-980 were fitted into the density map via MDFF. This

process is summarized in Figure S1; further details and validation of using MDFF for recovery of the all-atom resolution structure from an RBCG simulation will be discussed elsewhere.

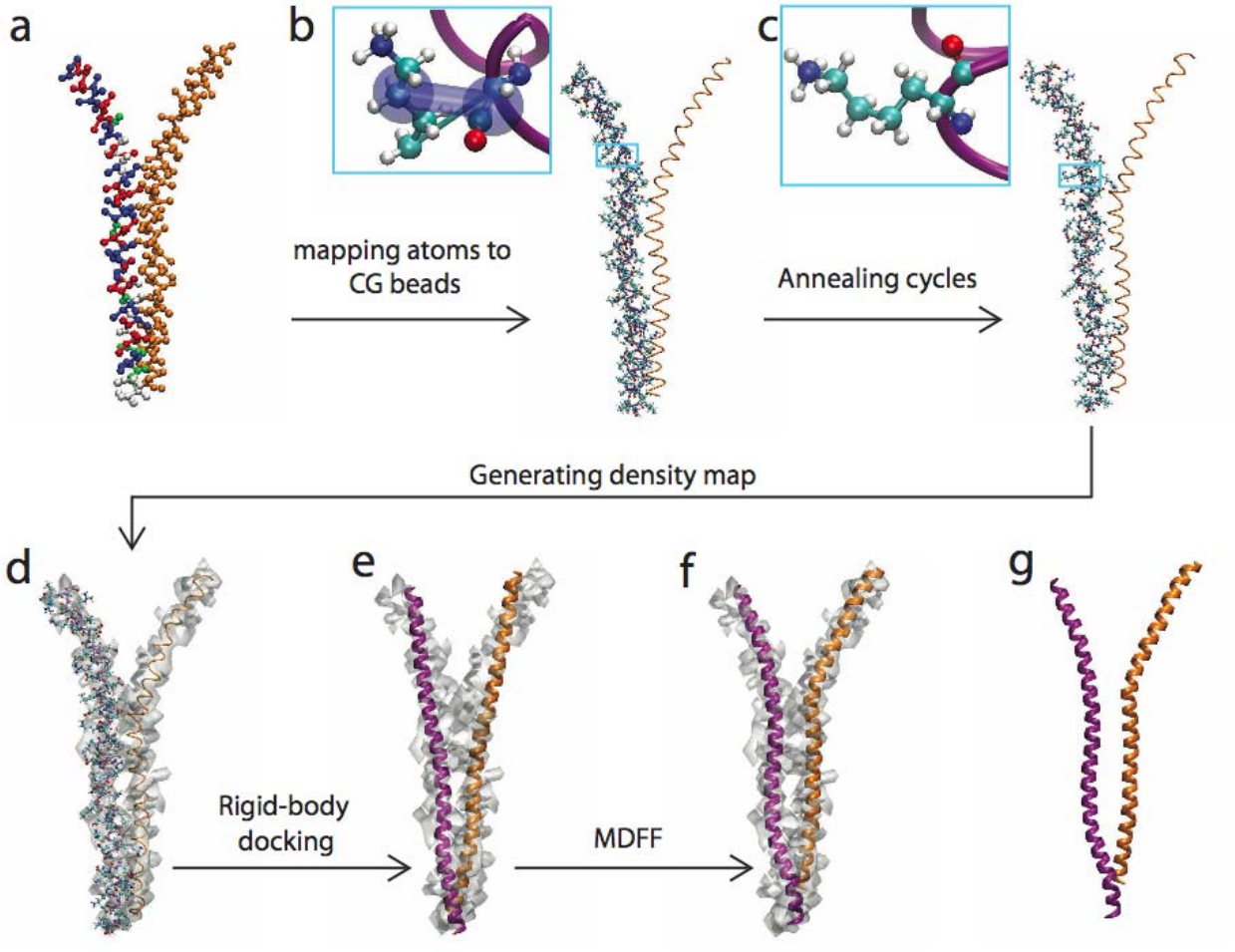


**Figure S1 (related to Figure 2).**
Recovery of all-atom resolution from a RBCG representation. A helix dimer in RBCG representation is shown in (a), with one helix shown in orange and the other colored by residue type (red: negatively charged; blue: positively charged; white: hydrophobic; green: polar). The RBCG beads are first converted to atoms by placing the center of mass position of a group of atoms to that of a corresponding bead, as shown in (b). Bond distortions in the all-atom representation can be seen after this step, displayed in the inset. The proper bond lengths can be regained through annealing cycles as shown in (c). However, secondary structure of the protein is unlikely to recover fully this way. The annealed structure in (c) is then used to generate an artificial electron density map, shown in (d). Two all-atom protein segments, shown in purple and orange in (e), with the correct secondary (i.e., α-helical) structure are placed in the density map by rigid-body docking, and then MDFF is applied to steer the protein into the map, as shown in (f). The resulting structure, displayed in (g), has the correct secondary structure characteristics, while also matching the conformation defined by the RBCG representation in (a).

**Equilibrium molecular dynamics**
After the all-atom resolution was recovered, the protein-solvent systems with MT helices, involving altogether over 155,000 atoms, were subjected to equilibrium (no constraints enforced)

all-atom MD using NAMD (Phillips *et al*. 2005. cited in the main text) with the CHARMM27 force field [4], including CMAP corrections [5]. Water molecules were described with the TIP3P model [6]. Long-range electrostatic forces were evaluated by means of the particle-mesh Ewald (PME) summation approach with a grid spacing of < 1Å. An integration time step of 2 fs was used in the framework of the Verlet r-RESPA algorithm [7]. Bonded terms and short-range, non-bonded terms were evaluated every time step, and long-range electrostatics every other time step. Constant temperature ($T$ = 310 K) was maintained using Langevin dynamics [8], with a damping coefficient of 1.0 $ps^{-1}$. A constant pressure of 1 atm was enforced using the Langevin piston algorithm [9] with a decay period of 200 fs and time constant of 50 fs. The methodology is reviewed in Phillips et al. (2005. cited in the main text).

**Free-energy calculation**

While equilibrium MD simulations of dimerized MT helices provide identification of molecular interactions contributing to the association of the helices, they do not quantify the energetic characteristics of the dimerization. To estimate the dimerization free energy of the MT helices, the adaptive biasing force (ABF) method was employed (Darve *et al.* 2001, Henin *et al*. 2004 and 2005. all cited in the main text; [10]), with the statistical error computed using the expression in Rodriguez-Gomez *et al*. (2004. cited in the main text). In an ABF calculation, instantaneous forces applied to the system in the direction of an a priori selected reaction coordinate were collected, and at intervals a biasing force is estimated and exerted to overcome local energy barriers. The free energy profile can then be recovered by thermodynamics integration. In the present study the reaction coordinate was chosen to be the center of mass separation of the two MT helices. The coordinate was probed over a 8-34 Å range, the interval being divided into several windows of size 3-5 Å for efficient sampling. For each window, at least 5 ns of MD trajectory was generated. ABF calculation was performed for the truncated, wild-type construct (Simulation ABF-2MT-907-to-940 in Table SS1), as well as the truncated and mutated construct (Simulation ABF-2MT-907-to-940-Mut in Table SS1).

Table S1. Simulations performed. In the "Type" column, "EQ-AA" denotes equilibrium all-atom simulations; "CG" denotes coarse-grained molecular dynamics; "Rev-CG" denotes reverting the CG representation into an all-atom representation; "ABF" denotes adaptive biasing free energy calculations.

| System | Type | Num. of Beads/Atoms | Simulation Time |
|---|---|---|---|
| Eq-1MT-907-to-980 | EQ-AA | 33,849 | 2 ns |
| CG-2MT-907-to-980 | CG | 3,120 | 1228 ns |
| Rev-CG-2MT-907-to-980 | Rev-CG | 155,197 | 12 ns |
| Eq-2MT-907-to-980 | EQ-AA | 40,344 | 50 ns |
| Eq-1MT-907-to-940 | EQ-AA | 14,314 | 2 ns |
| CG-2MT-907-to-940 | CG | 1,097 | 1016 ns |
| Rev-CG-2MT-907-to-940 | Rev-CG | 19,579 | 8 ns |
| Eq-2MT-907-to-940 | EQ-AA | 18,490 | 50 ns |
| FromFull-2MT-907-to-940-WT[1] | EQ-AA | 38,803 | 30 ns |
| FromFull-2MT-907-to-940-Mut[2] | EQ-AA | 38,786 | 30 ns |
| ABF-2MT-907-to-940 | ABF | 38,803 | 35 ns |

[1]Started from the end of Simulation Eq-2MT-907-to-980 but with only residues 907 to 940.
[2]Started from the end of Simulation Eq-2MT-907-to-980 but with only residues 907 to 940 and five hydrophobic residues mutated into glycines.

| | | | |
|---|---|---|---|
| ABF-2MT-907-to-940-Mut | ABF | 38,786 | 36 ns |

## SUPPLEMENTAL DATA

### Self-association of the myosin VI 907-940 construct

Park et al. (2006. cited in the main text) showed that the myosin VI 1-991 construct, but not the 1-917 construct, can dimerize, and later Mukherjea et al. (2009. cited in the main text) showed that the 1-940 construct also dimerizes, suggesting that there is a dimerization region in the MT domain of myosin VI. A series of simulations on the 907-940 segment were carried out in the present study to assess the dimerization propensity of this construct. For this purpose, a RBCG model of the 907-940 sequence was built, and the two segments were placed apart at the beginning of the simulation. After 200 ns, the segments were seen to associate and to maintain their association for the remainder of the 1-μs simulation. The dimerized segments were then reverse coarse grained into all-atom resolution as described above, with the all-atom system subsequently simulated for an additional 50 ns. Results of the simulation are summarized in Figure SS2.

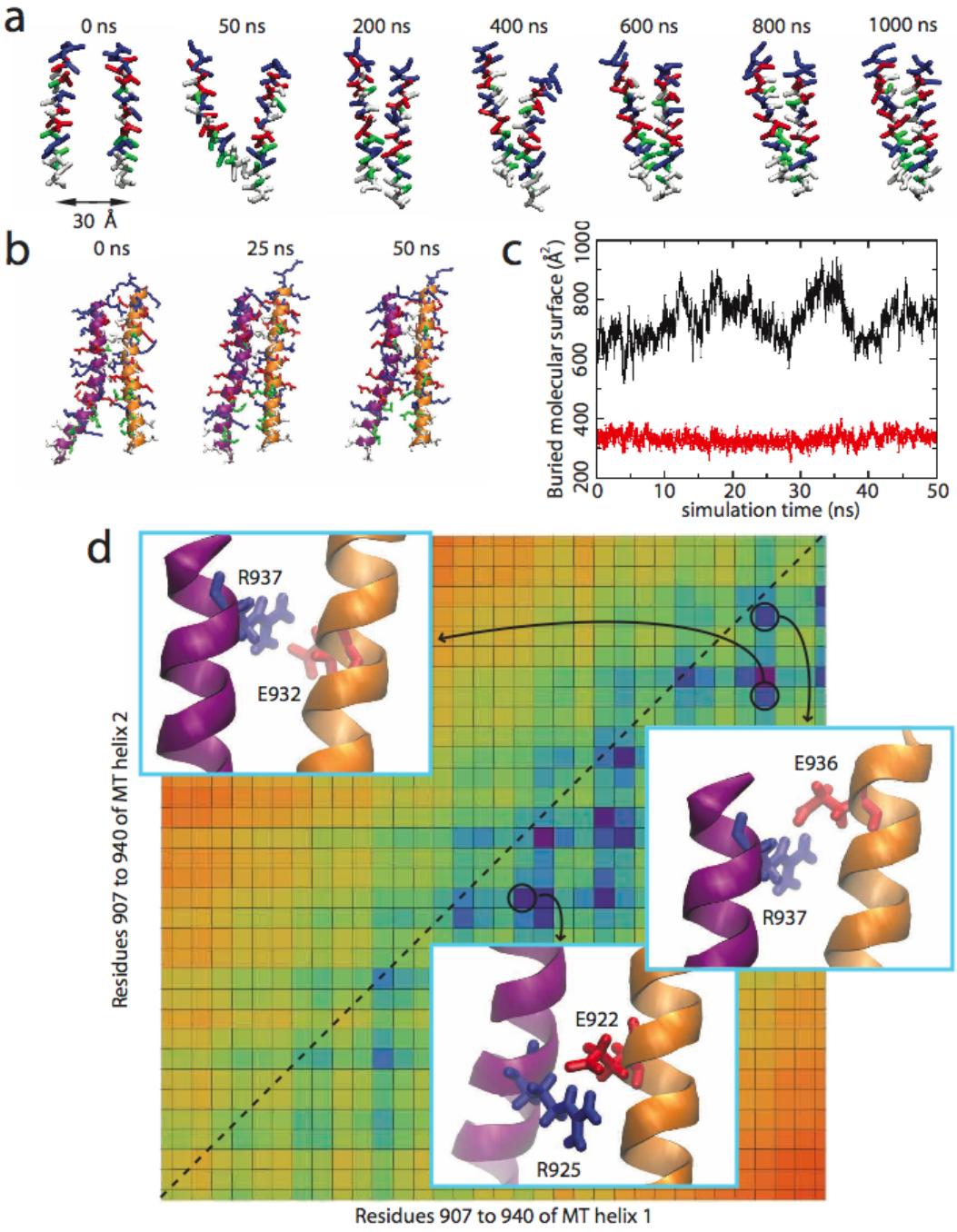


**Figure S2 (related to Figure 2).**
Dimerization of the truncated MT domain (907 to 940). (a) Spontaneous self-association of the MT segments in a 1-μs RBCG simulation. The MT segments are colored according to residue type. (b) All-atom MD of the MT dimer over a 50-ns duration. The dimer is seen to be stable with a significant buried molecular surface (black trace in (c)). The red trace in (c) is the buried molecular surface of the hydrophobic residues. The MT dimer contact map is shown in (d), with three observed salt-bridges displayed in the insets. The significant inter-helical interactions are seen to lie consistently below the *i=j* diagonal, denoted by a dashed line, indicating a vertical offset between the helices as also seen in the case of the 907-980 construct (Figure 2 of the main text).

**SUPPLEMENTAL REFERENCES**


[1] E. H. Lee, J. Hsin, M. Sotomayor, G. Comellas, and K. Schulten. Discovery through the computational microscope. *Structure*, 17:1295–1306, 2009.

[2] A. Y. Shih, P. L. Freddolino, S. G. Sligar, and K. Schulten. Disassembly of nanodiscs with cholate. *Nano Letters*, 7:1692–1696, 2007.

[3] W. Humphrey, A. Dalke, and K. Schulten. VMD – Visual Molecular Dynamics. *J. Mol. Graphics*, 14:33–38, 1996.

[4] A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, I. W. E. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B*, 102:3586–3616, 1998.

[5] A. D. MacKerell, Jr., M. Feig, and C. L. Brooks III. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry*, 25:1400–1415, 2004.

[6] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *Journal of Computational Chemistry*, 79:926–935, 1983.

[7] M. E. Tuckerman, B. J. Berne, and G. J. Martyna. Reversible multiple time scale molecular dynamics. *J. Phys. Chem. B*, 97:1990–2001, 1992.

[8] A. T. Brunger, C. L. Brooks III, and M. Karplus. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem. Phys. Lett.*, 105(5):495–498, March 1984.

[9] S. E. Feller, Y. H. Zhang, R. W. Pastor, and B. R. Brooks. Constant pressure molecular dynamics simulations — the langevin piston method. *J. Chem. Phys.*, 103:4613–4621, 1995.

[10] E. Darve, D. Wilson, and A. Pohorille. Calculating free energies using a scaled-force molecular dynamics algorithm. *Mol. Sim.*, 28:113–144, 2002.