Supporting Information

Materials and Methods

Materials. The coding sequence of Eg5 motor domain was first cloned into a baculovirus expression vector (Bacmid) and expressed in Sf9 insect cells**.** The Eg5 protein was then purified with affinity and gel filtration chromatography and assayed for its ATPase activity by monitoring the ADP production using the ADP Hunter reagents (from DiscoveRx). The *monastrol* compound was obtained from Tocris Bioscience.

Molecular Modeling. Based on the correlation-analysis of molecular dynamics simulation results, we have previously identified several novel allosteric sites (S1 to S4) on Eg5.[1] Since S3 and S4 are located in highly flexible loop regions, they were not considered in this study. The other two identified sites, S1 and S2, which are most consistent with reported mutagenesis results, along with the *monastrol*- and ADP-binding sites, were selected in this study to evaluate the STD-NMR/Docking/CORCEMA-ST method. Eg5 structural model was constructed based on the crystal structure of *monastrol*-Eg5 complex (PDB ID: 1Q0B).[2] The *monastrol* molecule was docked separately into the S1, S2, *monastrol*- and ADP-binding sites following an induced-fit docking (IFD) protocol,[3] which takes into consideration the ligand-induced receptor conformational changes in the binding pocket. Residues within 5 Å from the ligand were allowed to be flexible. The docked results were finally scored using the extra-precision (XP) scoring function of *Glide*. For each binding site, two best scored results with different docked conformations were selected as starting structures for the CORCEMA-ST calculations. All molecular modeling was performed using the *Maestro 9.3* suite from Schrödinger®.[4]

NMR Spectroscopy. The STD-NMR data were collected following established protocols.[5,6] Samples containing *monastrol* and Eg5 protein at ratios of 20:1 were prepared in D2O. STD-NMR spectra were recorded with a total of 32 K points, 80 scans, and selective saturation of protein resonances at 0, 0.65, 1.67, and 7.61 ppm (-8.18 ppm for the reference spectra), using a series of SEDUCE pulses (1000 points, 50 ms), for a total saturation time of 10 s (SEDUCE-1 pulse is similar to a Gaussian pulse, and has been used by other laboratories[7]). A saturation time of 10s was chosen to record steady-state STD intensities because of the long T1 values of some of the hydrogens on the ligand. Reference experiments using the free ligands themselves (i.e. without Eg5) were performed under the same experimental conditions to verify true ligand binding. No STD signals were present in the difference spectra of the free ligand, indicating that the effects observed in the presence of the protein were due to a true saturation transfer from the protein. ^{13}C T_1 NMR data for the free ligand were collected to estimate the free ligand correlation time of *monastrol*. The calculated T1 values of *monastrol*'s carbon-atoms with one directly-bonded proton were in the range of 400 to 600 ms, corresponding to rotational correlation times of 0.08 to 0.13 ns.

CORCEMA-ST Calculations. The CORCEMA-ST theory and the details of executing CORCEMA-ST has been described previously.[5, 6, 8] CORCEMA-ST program calculates the predicted STD-NMR intensities for any proposed molecular model of a ligand-receptor complex using parameters such as the correlation times, knowledge of saturated protein protons, exchange rates, and spectrometer frequency. The STD intensities were calculated as percentage fractional intensity changes $([(\boldsymbol{I}_{0(k)} - \boldsymbol{I}_{t(k)})*100]/\boldsymbol{I}_{0(k)}$, where k is a particular proton in the complex, and $\boldsymbol{I}_{0(k)}$ is its thermal equilibrium value from the intensity matrix $\boldsymbol{I}_{t(k)}$ and compared to the experimental STD values using an NOE *R*-factor defined as:[9]

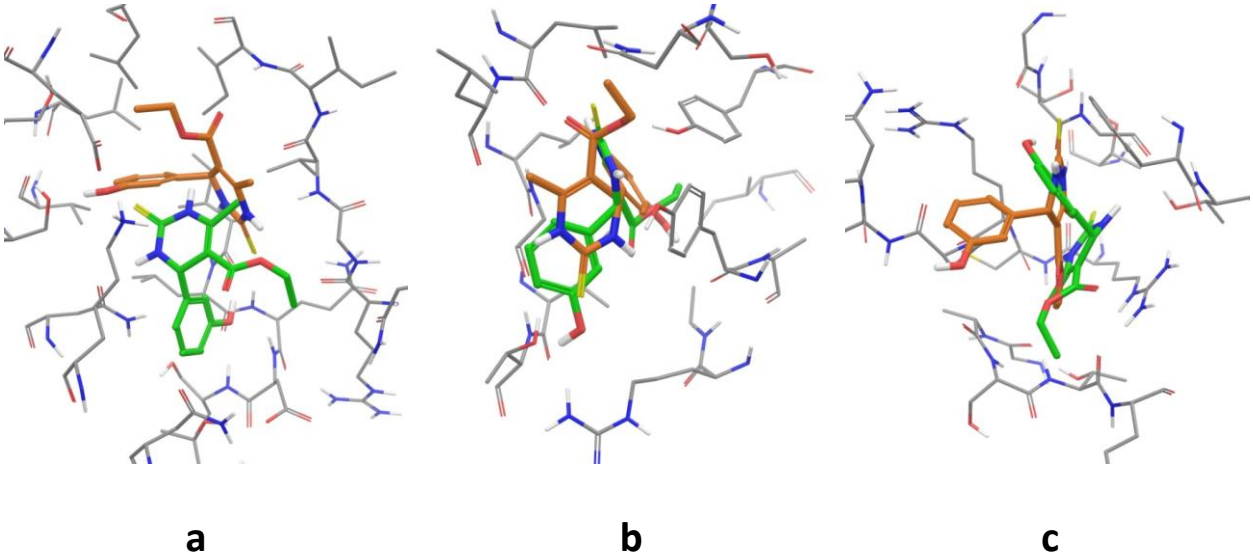
$$\text{NOE } R\text{-factor} = \sqrt{\frac{\sum(S_{exp,k}-S_{cal,k})^2}{\sum(S_{exp,k})^2}}$$

where, $S_{exp,k}$ and $S_{cal,k}$ refer to experimental and calculated STD values for proton k.

In our calculation, either the crystal structure (PDB ID: 1Q0B) or the docking generated *monastrol*-Eg5 complex structure was used as the staring model. Only residues that are within 6 Å from the ligand were considered in the matrix calculations. On the basis of our experimental conditions, the concentration of ligand was set as 500 µM and the ligand/protein ratio was kept fixed at 20:1. Since the purchased *monastrol* sample is a 50:50 mixture of two enantiomers (*R*- & *S*-) and the *R-monastrol* binds to Eg5 more than 20 times weaker than the *S-monastrol*,[10] we made the reasonable assumption that all the observed STDs arise from the binding of the *S-monastrol*. Thus the CORCEMA-ST calculations were done using a ligand concentration of 250 µM corresponding to the *S-monastrol* isomer. For methyl groups, a value of 0.85 was used for the order parameter S^2 and 2 ps was used for the internal motion correlation time. The leakage factor of 0.08 s^{-1} and a k_{on} of 10^8 s^{-1} M^{-1} were used. Other parameters were set based on published or measured results and modified within ±20% variation ranges of the reference values to obtain the best R-factor values: K_{eq} of the *monastol*-Eg5 complex was reported[11] to be in the range of 10^4-10^6 and the final reference value used in our CORCEMA-ST calculations was 5×10^5; the correlation time of the free ligand was set based on our ^{13}C T_1-derived values of 0.08 to 0.13 ns for different

carbons and the final reference value used in calculation was 0.1 ns; the correlation time of bound ligand was adjusted according to the published rotational correlation time of the entire Eg5 motor (~50 ns)[12] and the final reference value used was 50 ns.

Figure S1. Structural representation of docked models at a) S1 site; b) S2 site; and c) ATP-binding site. For each site, the two best scored docked-models used in CORCEMA-ST calculations were shown. The docked *monastrol* molecules were shown in solid sticks and colored in green (model-a) and orange (model-b), respectively. Eg5 residues within 4 Å of the docked ligands were shown in gray-colored tubes.



[1] W. Zhang, *J Phys Chem B* 2011, *115*(5), 784-795.
[2] Y. W. Yan, V. Sardana, B. Xu, C. Homnick, W. Halczenko, C. A. Buser, M. Schaber, G. D. Hartman, H. E. Huber, L. C. Kuo, *Journal of Molecular Biology* 2004, *335*(2), 547-554.
[3] W. Sherman, T. Day, M. P. Jacobson, R. A. Friesner, R. Farid, *J Med Chem* 2006, *49*(2), 534-553.
[4] *Maestro, version 9.0, Schrödinger, LLC, New York, NY*, 2009.
[5] V. Jayalakshmi, N. R. Krishna, *Journal of Magnetic Resonance* 2004, *168*(1), 36-45.
[6] N. R. Krishna, V. Jayalakshmi, *Progress in Nuclear Magnetic Resonance Spectroscopy* 2006, *49*(1), 1-25.
[7] J. W. Peng, C. A. Lepre, J. Fejzo, N. Abdul-Manan, J. M. Moore, *Methods Enzymol* 2001, *338*, 202-230.
[8] V. Jayalakshmi, N. R. Krishna, *Journal of Magnetic Resonance* 2002, *155*(1), 106-118.
[9] N. R. Krishna, D. G. Agresti, J. D. Glickson, R. Walter, *Biophys J* 1978, *24*(3), 791-814.
[10] S. DeBonis, J. P. Simorre, I. Crevel, L. Lebeau, D. A. Skoufias, A. Blangy, C. Ebel, P. Gans, R. Cross, D. D. Hackney, R. H. Wade, F. Kozielski, *Biochemistry* 2003, *42*(2), 338-349.
[11] L. Luo, J. D. Carson, D. Dhanak, J. R. Jackson, P. S. Huang, Y. Lee, R. Sakowicz, R. A. Copeland, *Biochemistry* 2004, *43*(48), 15258-15266.
[12] Z. Maliga, J. Xing, H. Cheung, L. J. Juszczak, J. M. Friedman, S. S. Rosenfeld, *Journal of Biological Chemistry* 2006, *281*(12), 7977-7982.