Supplementary Material for

Divergent allostery reveals critical differences between structurally homologous regulatory domains of *Plasmodium falciparum* and human protein kinase G

Jung Ah Byun[1], Bryan VanSchouwen[2], Jinfeng Huang[2], Ubaidullah Baryar[2], Giuseppe Melacini[1,2,*]

[1]Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, ON L8S 4M1, Canada.

[2]Department of Chemistry and Chemical Biology, McMaster University, Hamilton, ON L8S 4M1, Canada.

To whom correspondence should be addressed: melacin@mcmaster.ca

Supplementary Figures

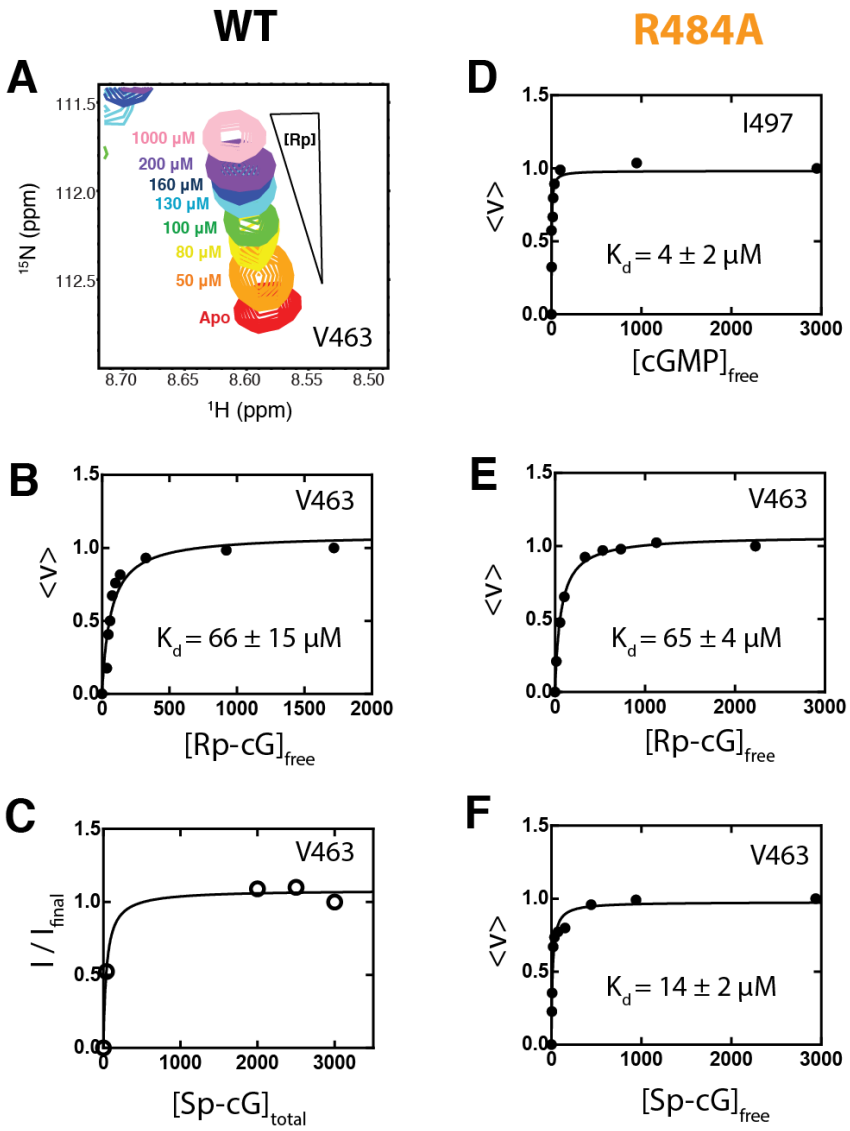

Figure S1. Binding isotherms and computed dissociation constants of cGMP, Rp-cGMPS and Sp-cGMPS for *pf*D WT and R484A. (A) ^{15}N-^{1}H HSQC-monitored titration of Rp-cG to WT *pf*D, zoomed into the peak for residue V463. **(B)** Binding isotherm of Rp-cG titration to WT *pf*D, showing the fraction of bound Rp-cG (*i.e.* <v>) *versus* the concentration of unbound Rp-cG (*i.e.* [Rp-cG]$_{free}$). <v> was computed using the NMR-monitored titration shown in panel (A) and the associated chemical shift changes. **(C)** Binding of Sp-cG to WT *pf*D is slow relative to the NMR timescale, hence concentration-dependent chemical shift changes were not observed. Therefore, intensity ratios of the peak for residue V463, relative to the intensity of the peak from the last titration point (*i.e.* 3mM Sp-cG), were used instead. **(D)** Binding isotherm of cGMP to R484A *pf*D. **(E)** Similar to panel (D), but for Rp-cG. **(F)** Similar to panel (D), but for Sp-cG.

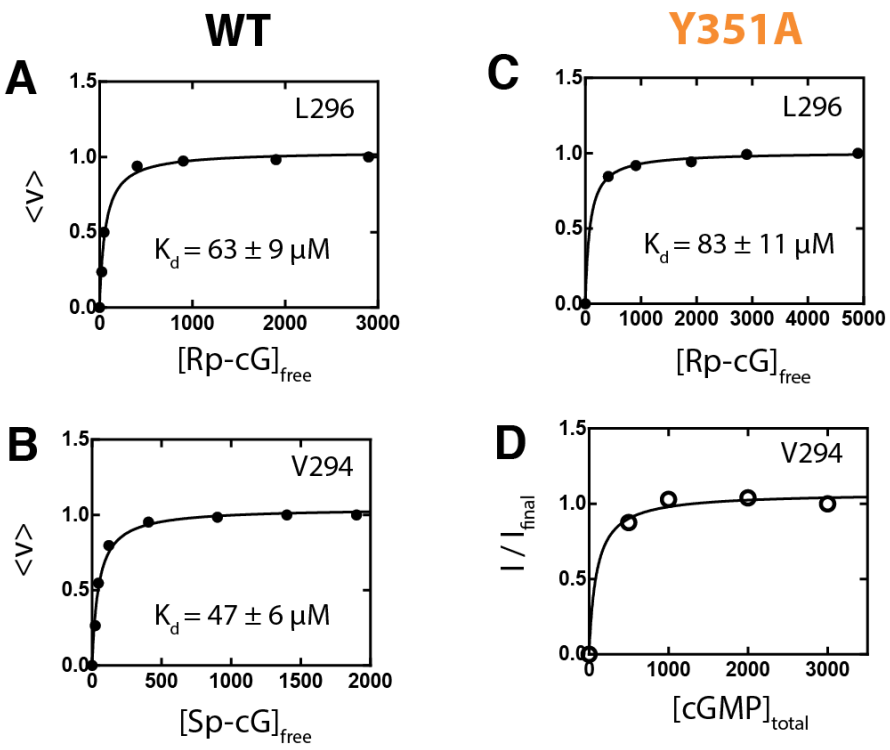

Figure S2. Binding isotherms of cGMP, Rp-cGMPS and Sp-cGMPS for *h*B WT and Y351A. (A) Binding isotherm of Rp-cG titration to WT *h*B, showing the fraction of bound Rp-cG (*i.e.* <v>) *versus* the concentration of unbound Rp-cG (*i.e.* $[\text{Rp-cG}]_{free}$). <v> was computed using an NMR-monitored titration and the associated chemical shift changes. **(B)** Similar to panel (A), but for Sp-cG. **(C)** Similar to panel (A), but for the Y351A mutant. **(D)** Binding of cGMP to *h*B is slow relative to the NMR timescale, hence concentration-dependent chemical shift changes were not observed. Therefore, intensity ratios of the peak for residue V294, relative to the intensity of the peak from the last titration point (*i.e.* 3mM cGMP), were used instead.

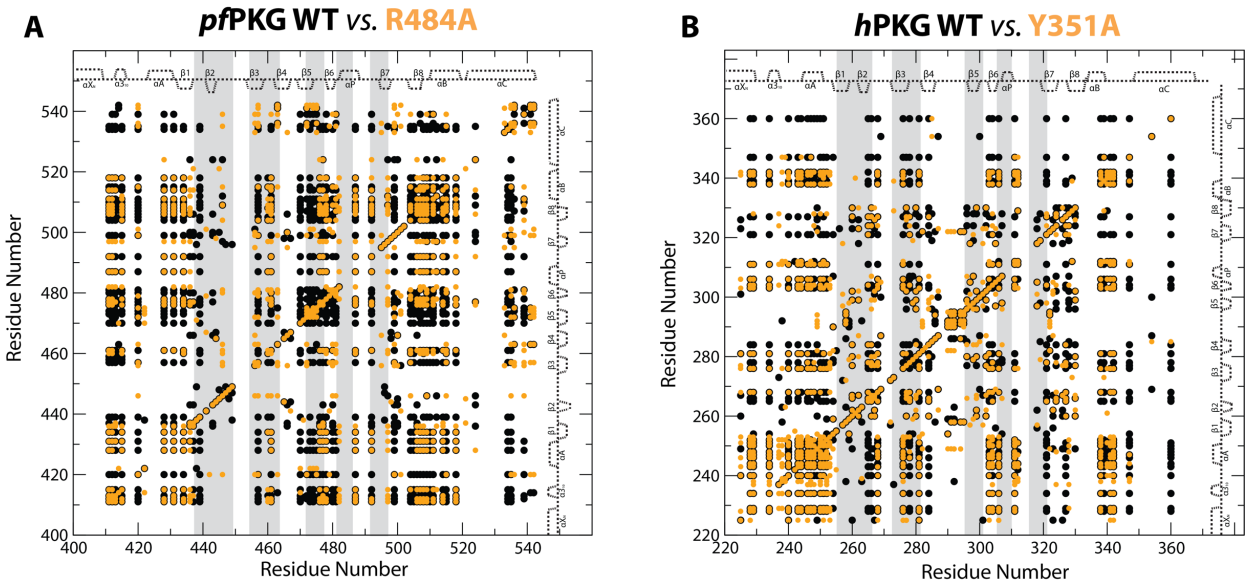

Figure S3. CHESCA correlation matrices for WT and related capping residue mutants of (A) *pf*PKG and (B) *h*PKG as in Figure 3 but with regions of homologous residues highlighted in grey. The cut-off value for the Pearson correlation coefficient was set to 0.95.

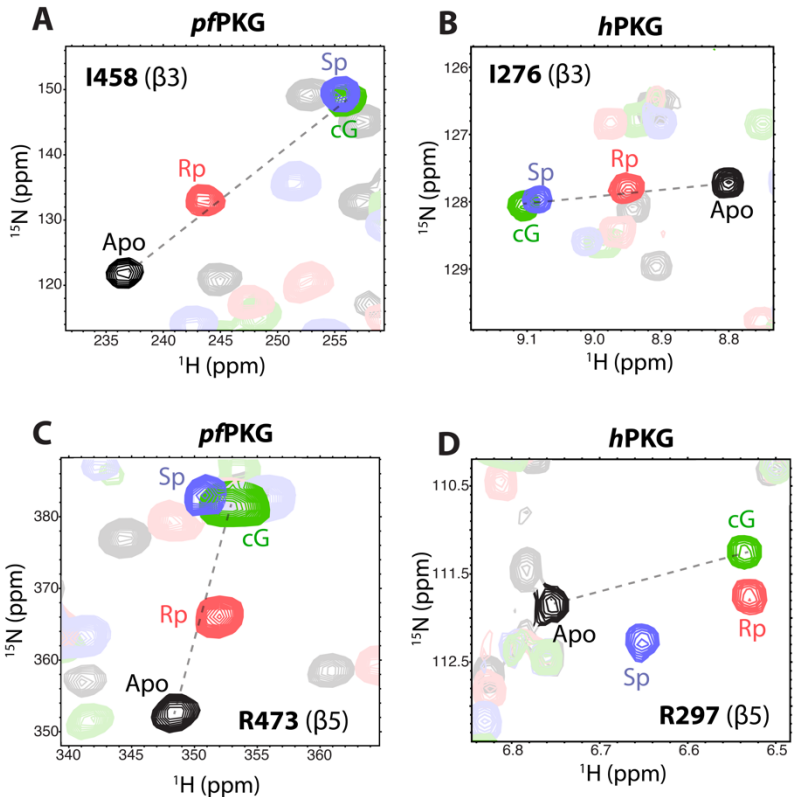

Figure S4. Spectral comparisons of homologous residues in WT *pf*D and *h*B. Overlay of ^{15}N-^{1}H HSQC spectra of apo (black), Rp-cGMPS-bound (red), Sp-cGMPS-bound (purple) and cGMP-bound (green) zoomed into I458 of *pf*D (A) and its homologous residue in *h*B, I276 (B). The relative positions of HSQC cross-peaks for cGMP-analog bound states are similar for these homologous residues. (C) Similar to panel (A), but for R473 of *pf*D. (D) Similar to panel (C), but for its homologous residue in *h*B, R297. The relative positions of the HSQC cross-peaks are now different for these homologous residues.

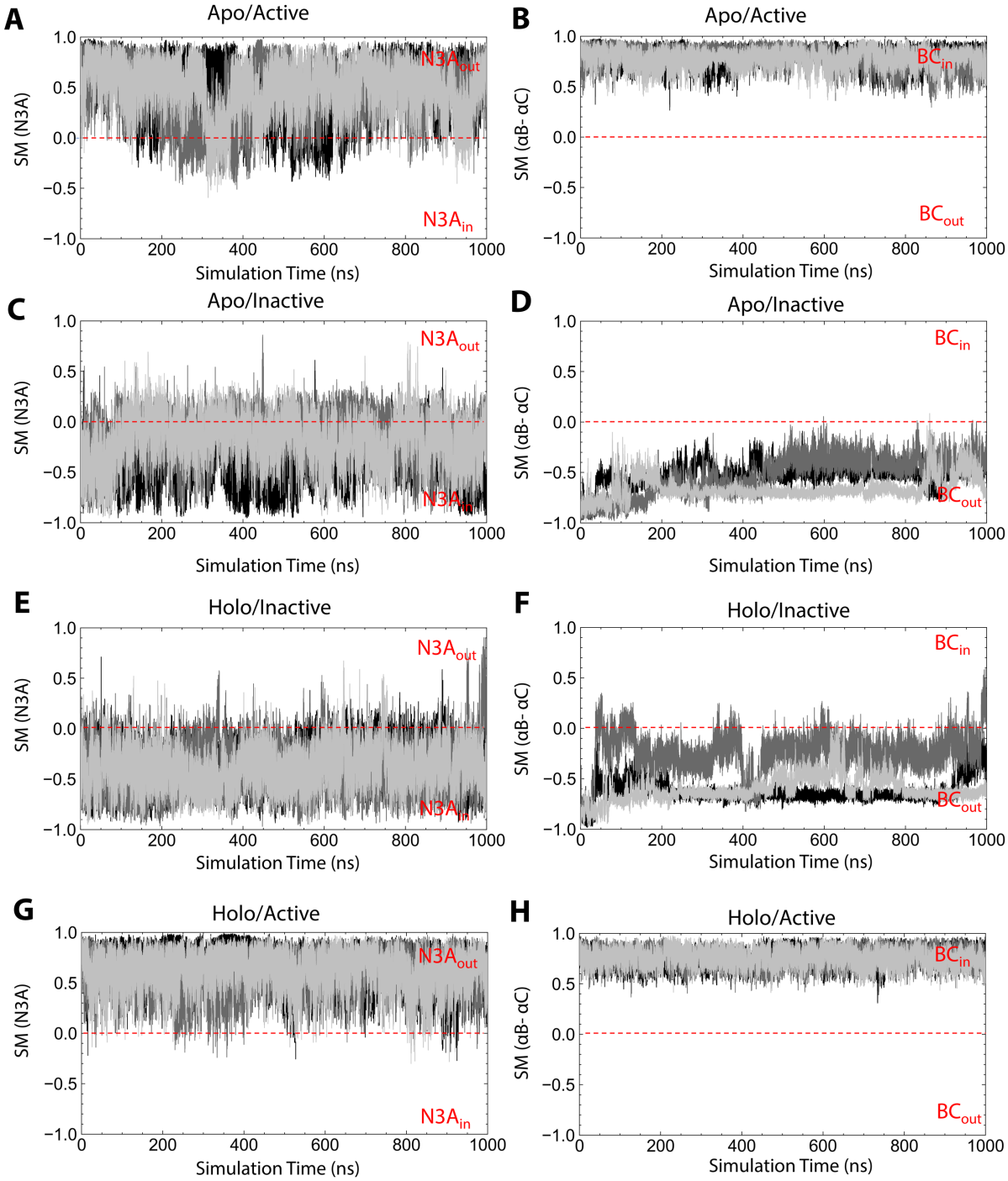

Figure S5. Similarity measurements (SM) for *h*B over time. (A, C, E, G) Similarity measurements over time for the N3A region of the four states in Fig. 1H. An SM value close to 1 indicates similarity to the "out" conformation (holo/active) while -1 captures the "in" conformation (apo/inactive). **(B, D, F, H)** Similarity measurements over time for the αB-αC region of the four states in Fig. 1H. If the SM value is close to 1, the αB-αC helices approach the "in" conformation (holo/active), while an SM value of -1 captures the "out" conformation (apo/inactive).

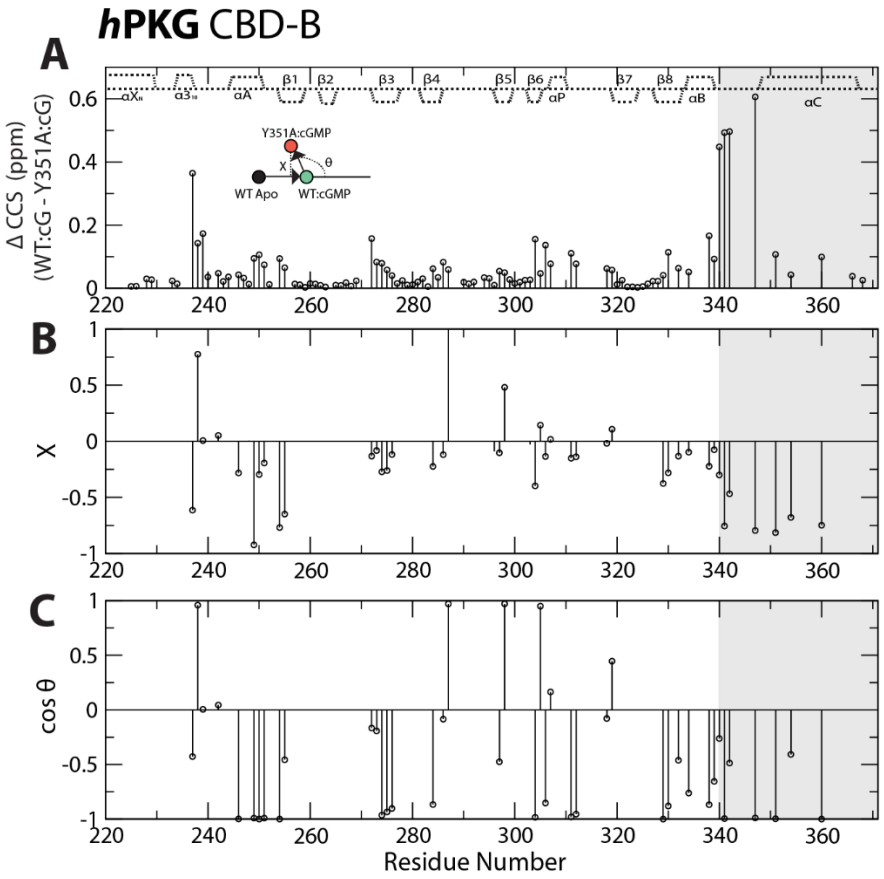

Figure S6. Chemical Shift Projection Analysis (CHESPA) of *h*PKG CBD-B Y351A mutant. **(A)** Compounded ^{15}N-^{1}H chemical shift changes (ΔCCS) between cGMP-bound WT and cGMP-bound Y351A samples of *h*B, plotted *versus* residue number. The CHESPA vector scheme is shown on the left side of the plot. **(B)** CHESPA fractional shift (X) values *versus* residue number for cGMP-bound Y351A *h*B. **(C)** CHESPA cos(θ) values *versus* residue number for cGMP-bound Y351A *h*B.

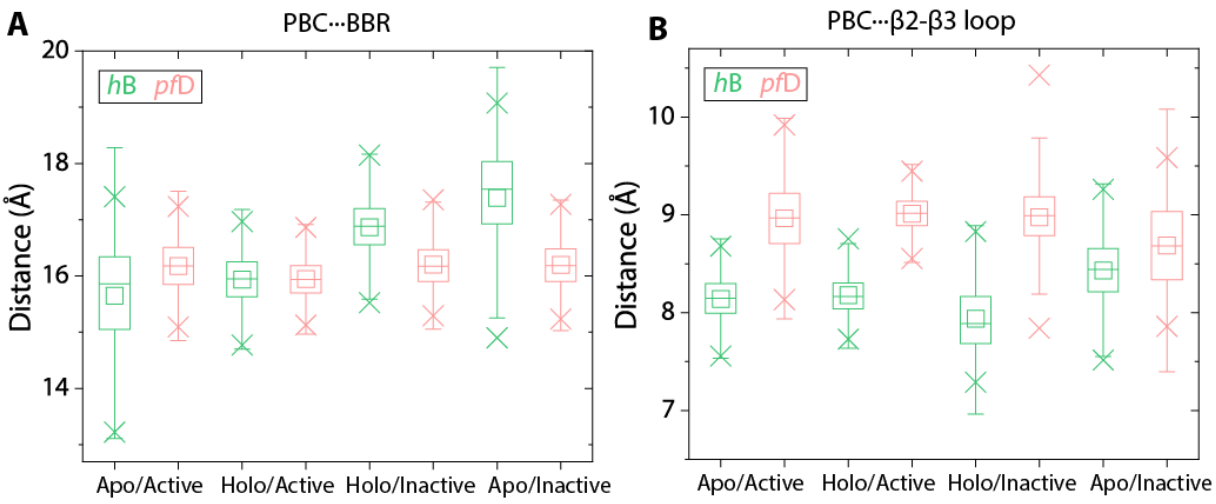

Figure S7. Distances between the PBC and BBR/β2-3 loop, as computed from MD simulations of WT *h*B and *pf*D. For each state, three 1-μs MD trajectories were used for the analysis. (A) Boxplots for distances between the PBC and BBR of *h*B (green) and *pf*D (red)(1). **(B)** Boxplots for distances between the PBC and β2-3 loop of *h*B (green) and p*f*D (red)(1). In the boxplots, the middle, bottom, and top lines of the box define the median, 25th percentile and 75th percentile, respectively; the whiskers indicate additional data falling within 1.5*IQR (where "IQR" is the interquartile range) above the 75th percentile or below the 25th percentile; the small square represents the mean, and the two crosses represent the 1st and 99th percentiles of the data.

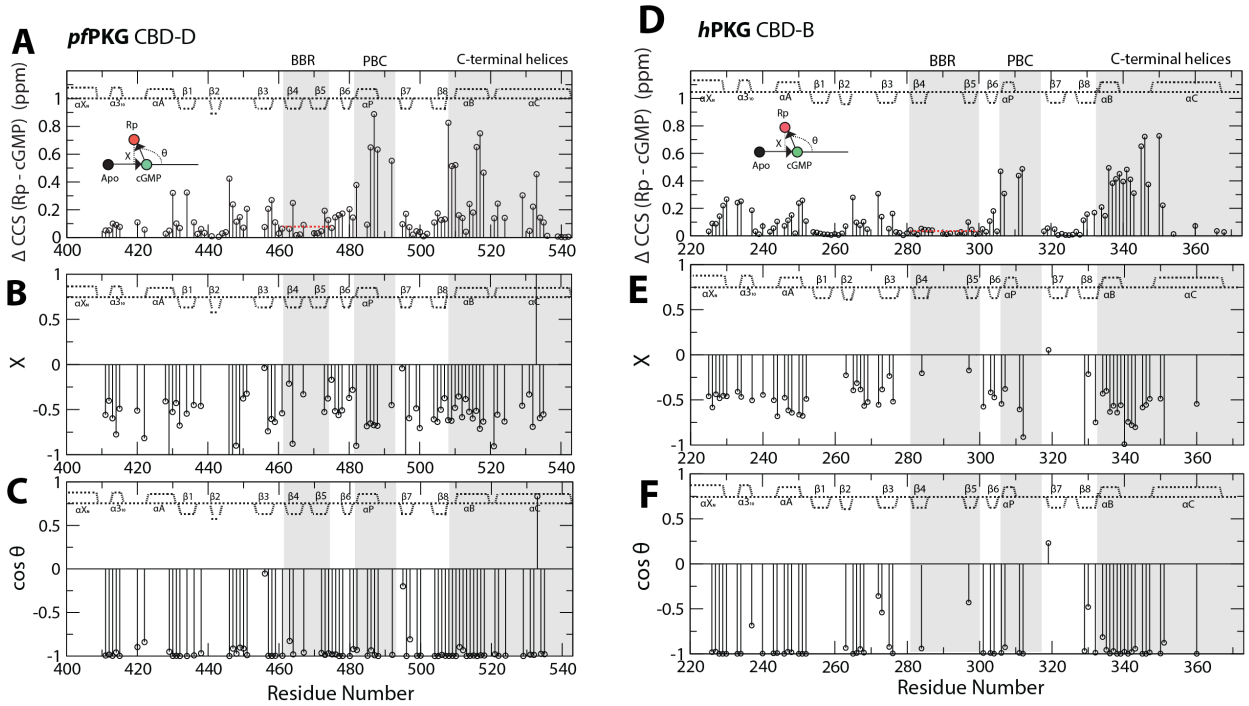

Figure S8. The *pf*PKG BBR exhibits significant perturbations compared to the *h*PKG BBR. (A) Compounded ^{15}N-^{1}H chemical shift changes (ΔCCS) between cGMP-bound and Rp-cGMPS-bound *pf*PKG CBD-D, plotted *versus* residue number. The red dashed line represents the average ΔCCS of residues in the BBR. The CHESPA vector scheme is shown on the left side of the plot. (**B)** CHESPA fractional shift (X) values *versus* residue number for Rp-cGMPS-bound *pf*D. **(C)** CHESPA cos(θ) values *versus* residue number for Rp-cGMPS-bound *pf*D. **(D-F)** Similar to panels (A-C), respectively, but for *h*B.

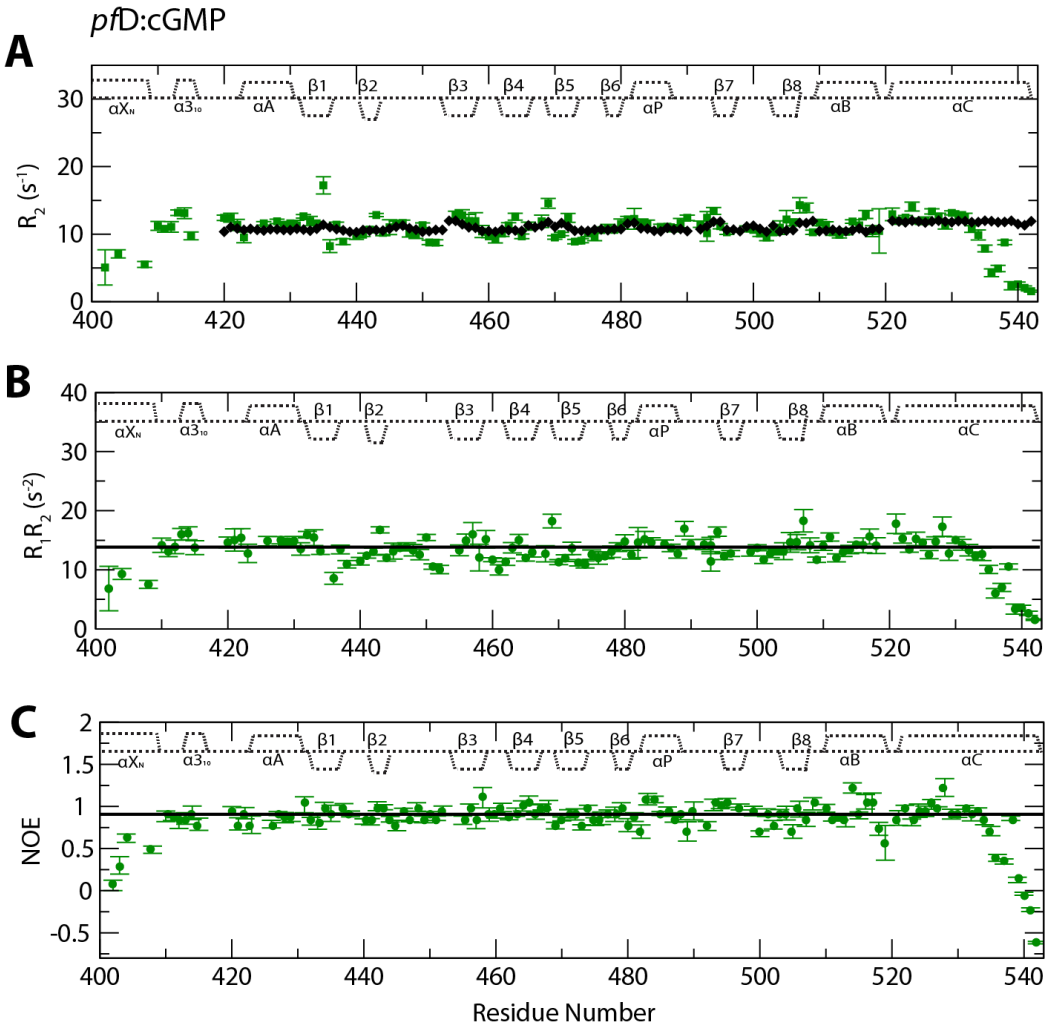

Figure S9. Backbone ^{15}N relaxation rates for cGMP-bound *pf*D. (A) R_2 relaxation rates. The black diamonds represent simulated R_2 rates computed with HydroNMR(2). **(B)** Product of R_2 and R_1 relaxation rates. **(C)** ^{1}H,^{15}N NOE values. Black horizontal lines denote the average values calculated for the rigid inner β-strands of the β-barrel, which are assumed to sense primarily the overall tumbling motion of the protein in solution.

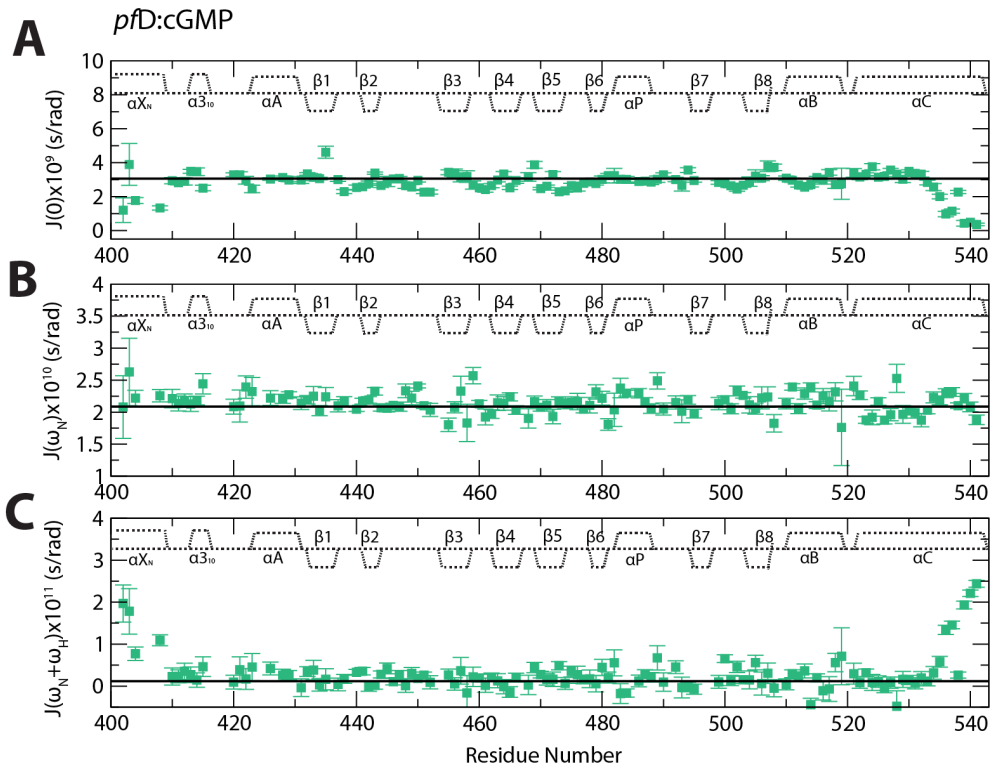

Figure S10. Reduced spectral densities for cGMP-bound *pf*D. (A) J(0) values, **(B)** J(ω_N) values, and **(C)** J($\omega_N + \omega_H$) values. Black horizontal lines denote the average values calculated for the rigid inner β-strands of the β-barrel, which are assumed to sense primarily the overall tumbling motion of the protein in solution.

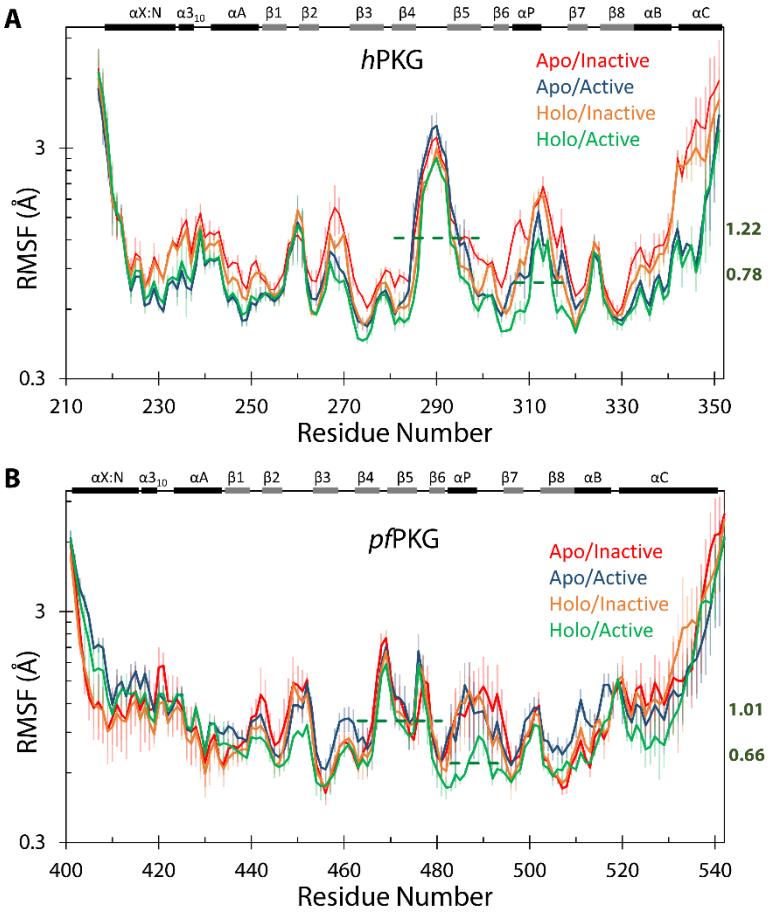

Figure S11. Residue-specific structural fluctuations, as computed from MD simulations of WT *h*B and *pf*D. (A) Root-mean-square fluctuations (RMSFs) *vs.* residue number for *h*PKG CBD-B, computed *via* overlay of the Cα atoms of the whole *h*PKG CBD-B domain to the respective initial models. Green dashed lines represent the average RMSFs for the BBR and PBC regions of the holo/active state, and the corresponding values are indicated along the right y-axis. A $\log_{10}$ scale is used for the RMSF values. **(B)** Similar to panel (A), but for *pf*PKG CBD-D which is derived from previous MD simulations(1).

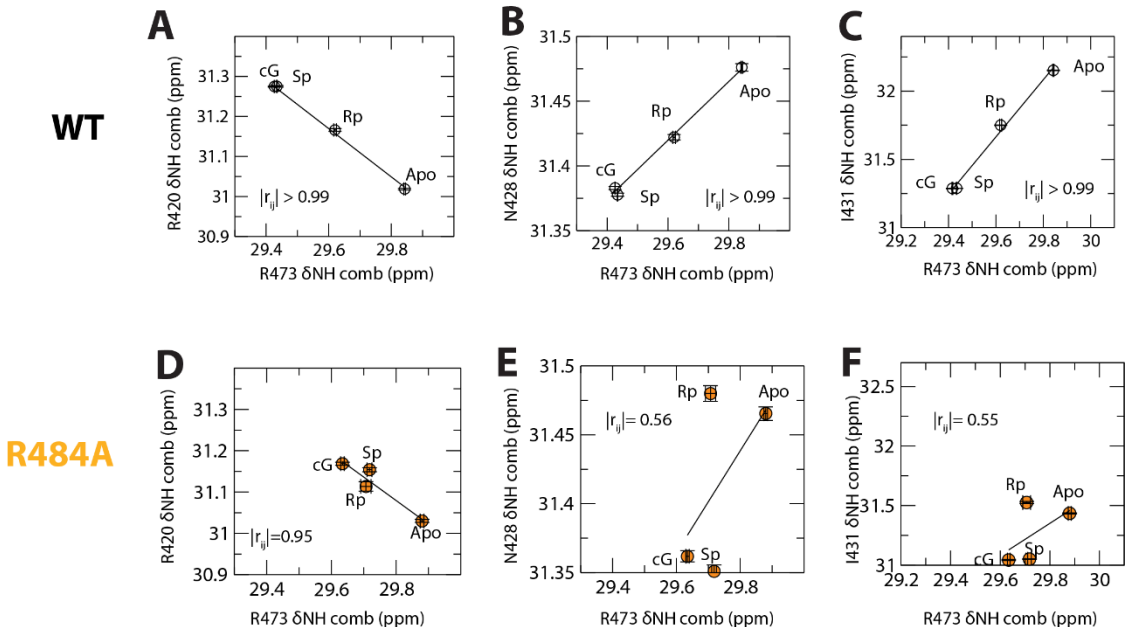

Figure S12. Pairwise inter-residue correlation plots of the guanine-interacting BBR residue R473 with residues in the N3A (R420, N428, I431) in **(A-C)** WT *pf*D and **(D-F)** R484A *pf*D.

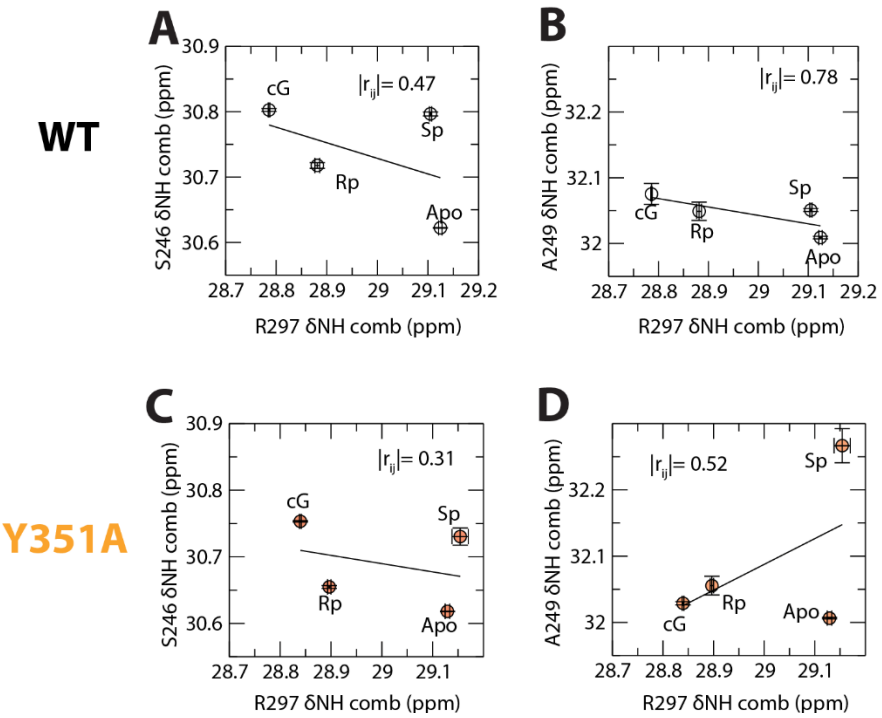

Figure S13. Pairwise inter-residue correlation plots of the guanine-interacting BBR residue R297 with residues in the N3A (S246 and A249) in **(A, B)** WT *h*B and **(C, D)** Y351A *h*B.

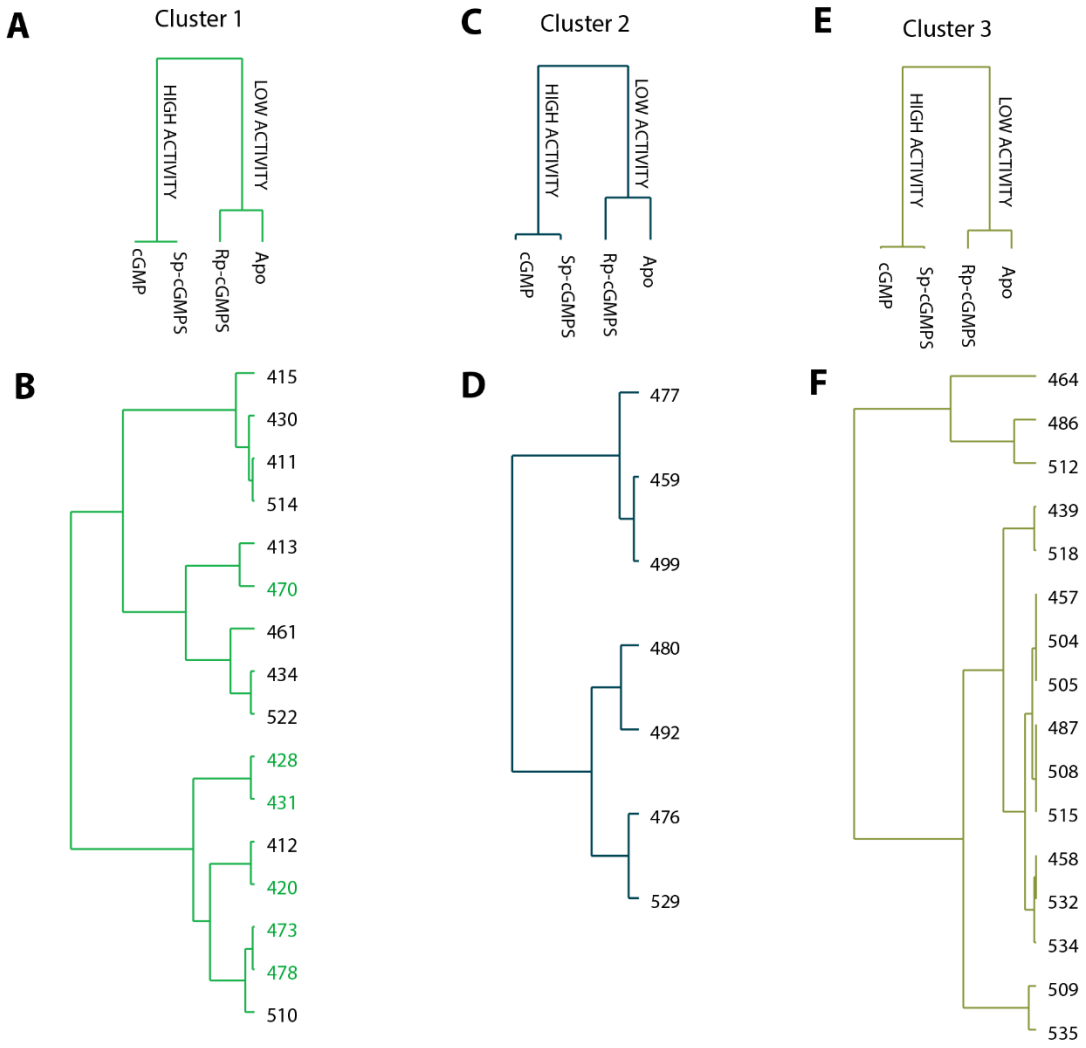

Figure S14. Dendrograms of allosteric residue clusters from WT *pf*D. (A) Dendrogram of the agglomerative clustering of the four CHESCA perturbation states (*i.e.* apo, cGMP- and cGMP-analog-bound states) using residues from cluster 1. **(B)** Dendrogram of the residues in cluster 1, which includes R473 and is the cluster used for mapping on the structure in Fig. 6I. Residue numbers highlighted in green indicate residues used for the pairwise correlation plots in Fig. 6A-D and S12. **(C)** Similar to panel (A) but for cluster 2. **(D)** Dendrogram of cluster 2 residues. **(E)** Similar to panel (A) but for cluster 3. **(F)** Dendrogram of cluster 3 residues.

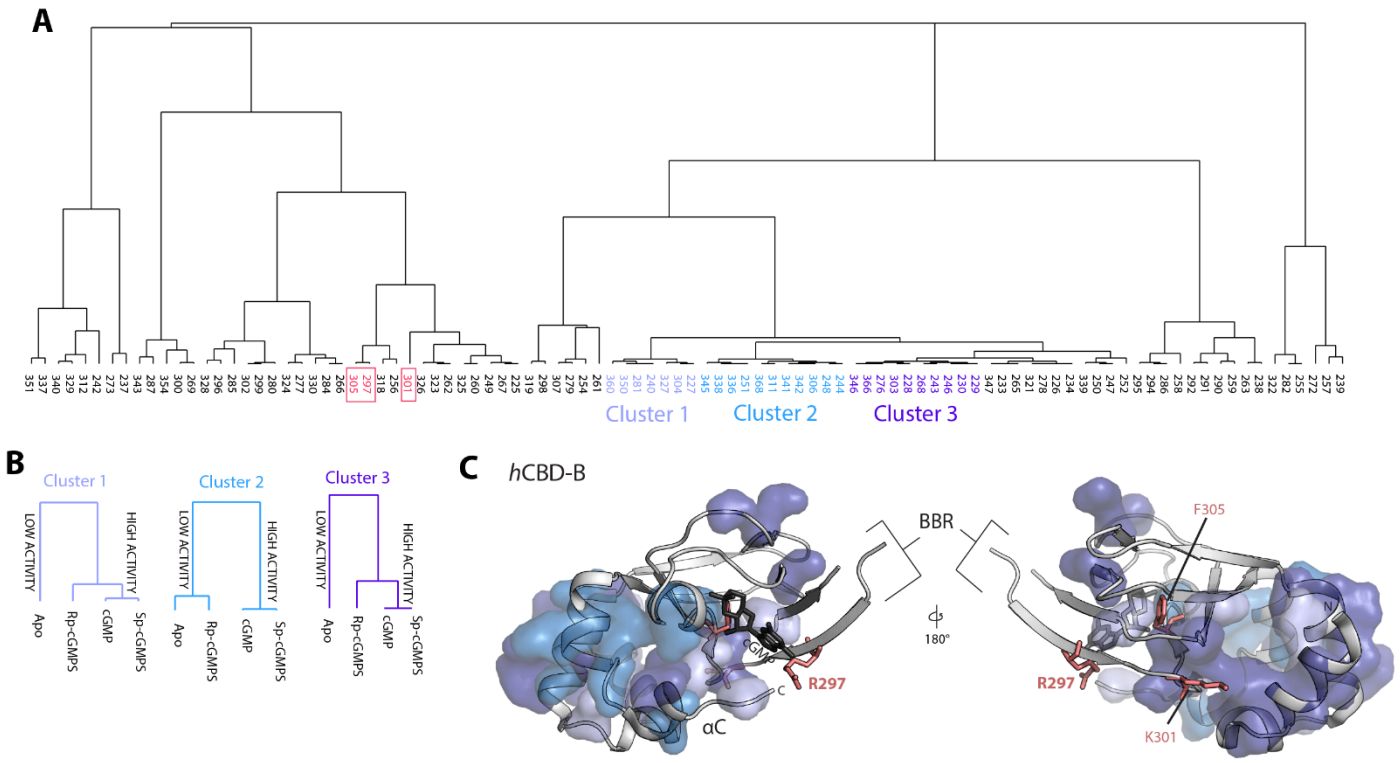

Figure S15. Agglomerative clustering analysis of WT *h*B. (A) Dendrogram of all residues from *WT hB*. Allosteric clusters containing residues that are significantly correlated (*i.e.* with a correlation coefficient > 0.95) are highlighted (*i.e.* clusters 1-3 indicated in the dendrogram). Residue numbers in red boxes indicate residues used for the pairwise correlation plots in Fig. 6E-H, including BBR residue R297. **(B)** Dendrogram of the agglomerative clustering of the four CHESCA perturbation states (*i.e.* apo, cGMP- and cGMP-analog-bound states) using residues from each of the three allosteric clusters in *hB*. **(C)** Map of the residues of the allosteric clusters defined within *h*B, shown as a surface representation. The residues from Fig. 6E-H are represented as pink sticks, and are not part of any of the allosteric clusters defined in *h*B.

References

1. Huang, J., Byun, J. A., VanSchouwen, B., Henning, P., Herberg, F. W., Kim, C., and Melacini, G. (2021) Dynamical Basis of Allosteric Activation for the Plasmodium falciparum Protein Kinase G. *J. Phys. Chem. B*. **125**, 6532–6542
2. García de la Torre, J., Huertas, M. ., and Carrasco, B. (2000) HYDRONMR: Prediction of NMR Relaxation of Globular Proteins from Atomic-Level Structures and Hydrodynamic Calculations. *J. Magn. Reson.* **147**, 138–146