Supporting Information

Ferreiro et al. 10.1073/pnas.1018980108

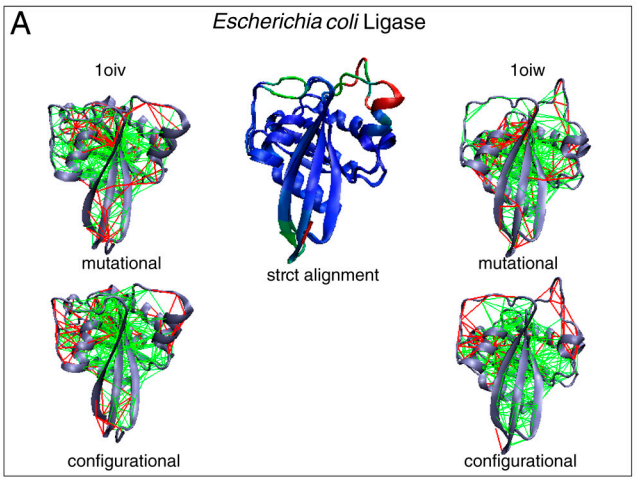
SI Methods

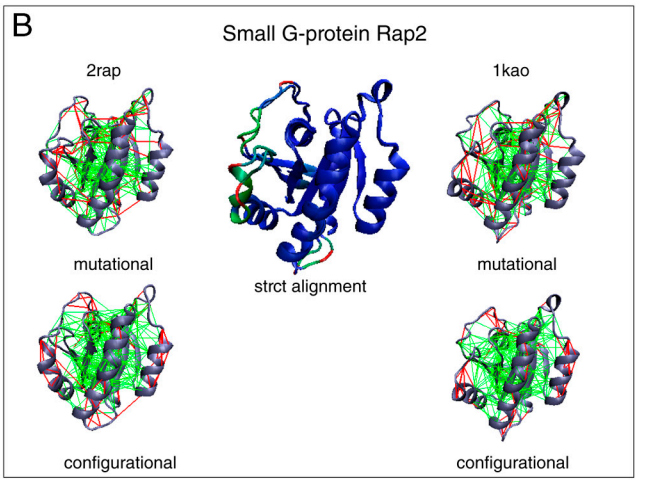
Local Frustration Definitions. Localizing energetic frustration requires the evaluation of the energy of a protein in its native state and comparison to the energies of a set of "decoy" states. The algorithm we use requires as input a high-resolution structure and an accurate energy function (1). We chose to base our energy function on the associative memory Hamiltonian optimized with water-mediated interactions (2, 3). A contact is defined as "minimally frustrated" if its native energy is at the lower end of the distribution of decoy energies, having a frustration index as measured with a Z score of 0.78 or higher magnitude, that is, the majority of other amino acid pair in that position would be unfavorable (1). Conversely, a contact is be defined as "highly frustrated" if the native energy is at the other end of the distribution with a local frustration index lower than −1, that is, unlike for a minimally frustrated pair, most other amino acid pairs at that location would be more favorable for folding than the native ones by more than one standard deviation of that distribution. If the native energy is in between these limits we define the contact as "neutral."

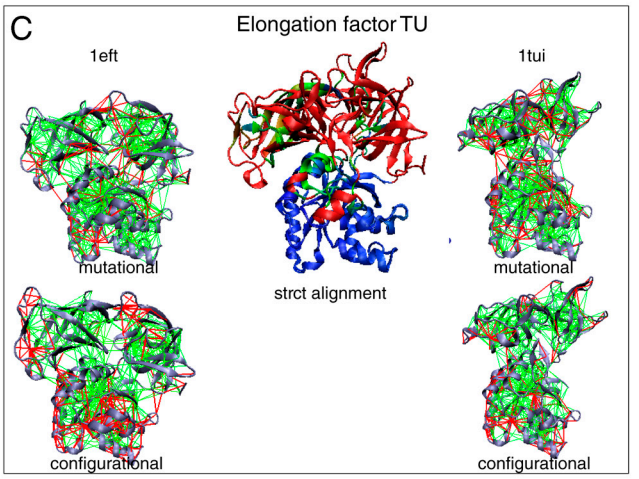
Projecting Local Frustration Information in Sequence Space. A virtual particle was defined at the geometrical center of each contact considering only the positions of the interacting Cα. The virtual particles within 5$\mathring{A}$ of a given Cα were counted and classified as minimally frustrated, neutral, or highly frustrated by their frustration index as described above.

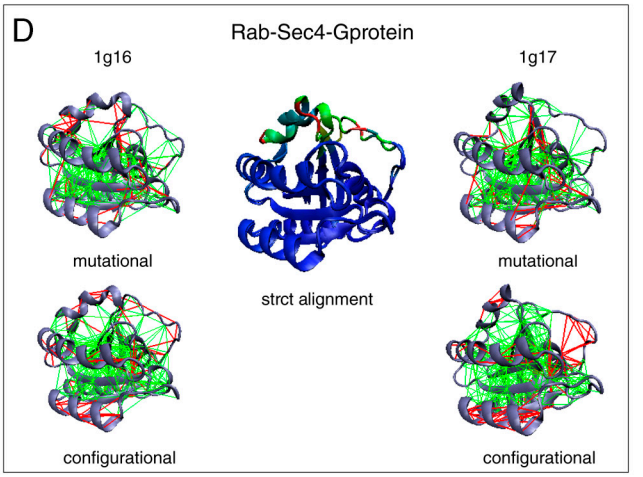
Visualization and Numerical Tools. All the visual representations of the proteins were done using the program VMD (4). The contacts were drawn between the Cα atoms of each amino acid. Pair distribution functions were calculated using Matlab (The MathWorks, Inc.) and the plots generated with ProFit (Quantum Soft).

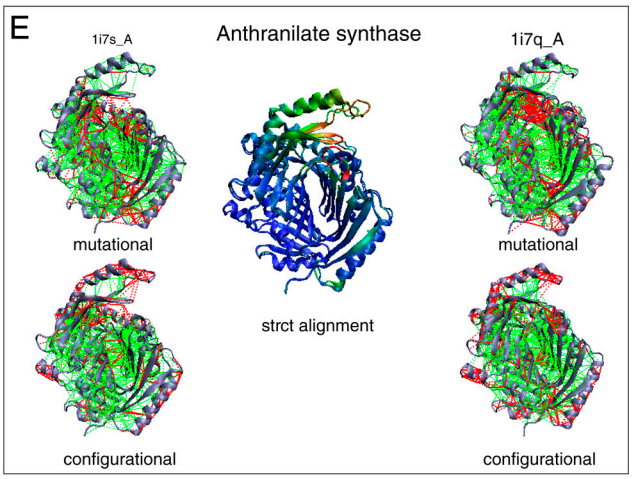
1. Ferreiro DU, Hegler JA, Komives EA, Wolynes PG (2007) Localizing frustration in native proteins and protein assemblies. *Proc Natl Acad Sci USA* 104:19819–19824.
2. Papoian GA, Ulander J, Wolynes PG (2003) Role of water mediated interactions in protein–protein recognition landscapes. *J Am Chem Soc* 125:9170–9178.
3. Papoian GA, Wolynes PG (2003) The physics and bioinformatics of binding and folding-an energy landscape perspective. *Biopolymers* 68:333–349.
4. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14:33–38.

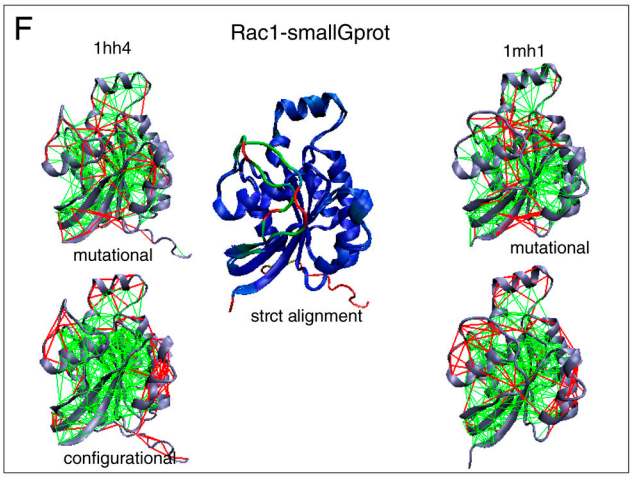

A
Escherichia coli Ligase
1oiv
1oiw
mutational
strct alignment
mutational
configurational
configurational

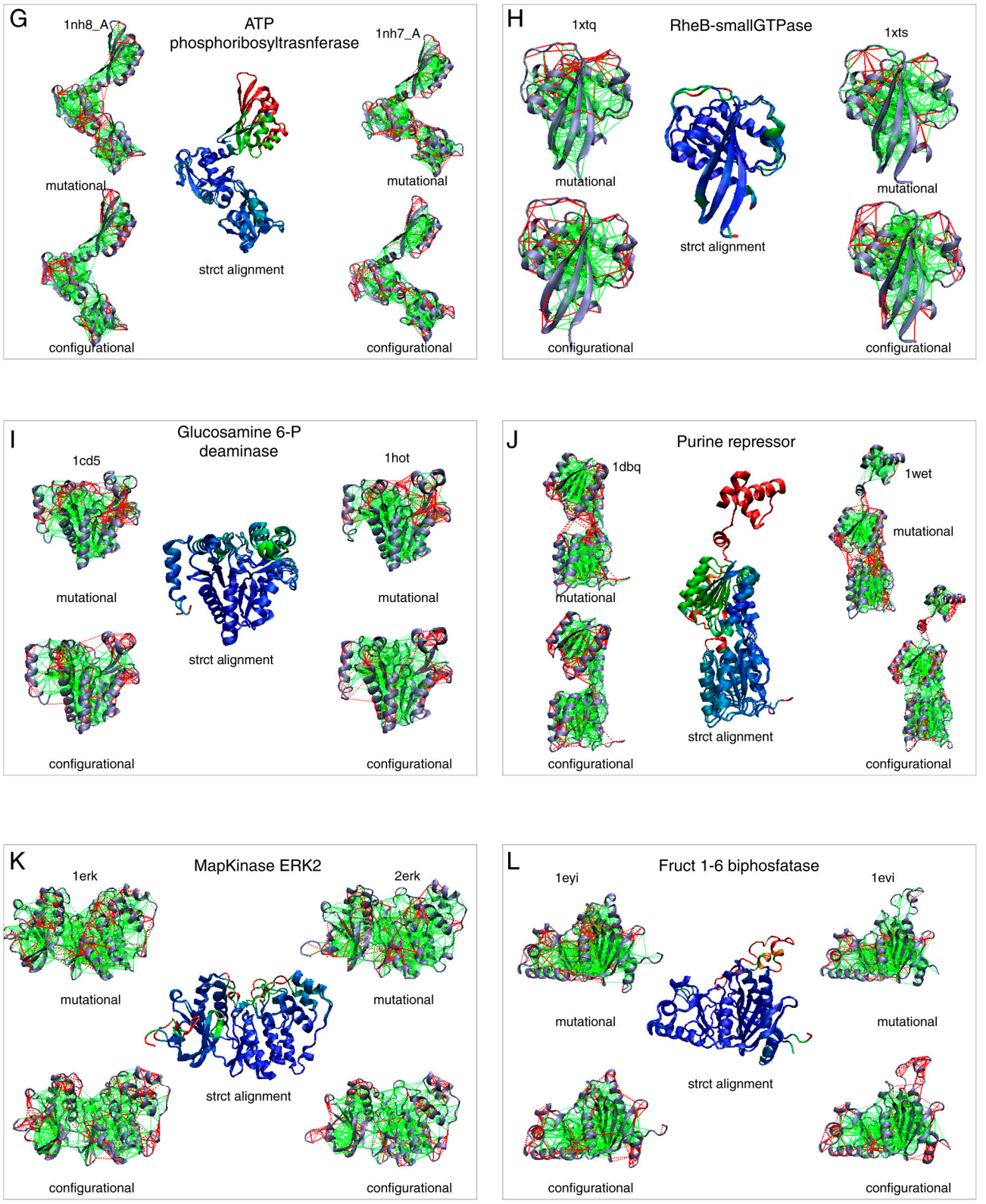

B
Small G-protein Rap2
2rap
1kao
mutational
strct alignment
mutational
configurational
configurational


C
Elongation factor TU
1eft
1tui
mutational
strct alignment
mutational
configurational
configurational


D
Rab-Sec4-Gprotein
1g16
1g17
mutational
strct alignment
mutational
configurational
configurational


E
Anthranilate synthase
1i7s_A
1i7q_A
mutational
strct alignment
mutational
configurational
configurational


F
Rac1-smallGprot
1hh4
1mh1
mutational
strct alignment
mutational
configurational


G
1nh8_A
ATP phosphoribosyltrasnferase
1nh7_A
mutational
mutational
strct alignment
configurational
configurational
H
1xtq
RheB-smallGTPase
1xts
mutational
mutational
strct alignment
configurational
configurational
I
Glucosamine 6-P deaminase
1cd5
1hot
mutational
mutational
strct alignment
configurational
configurational
J
Purine repressor
1dbq
1wet
mutational
mutational
strct alignment
configurational
configurational
K
MapKinase ERK2
1erk
2erk
mutational
mutational
strct alignment
configurational
configurational
L
Fruct 1-6 biphosfatase
1eyi
1evi
mutational
mutational
strct alignment
configurational
configurational

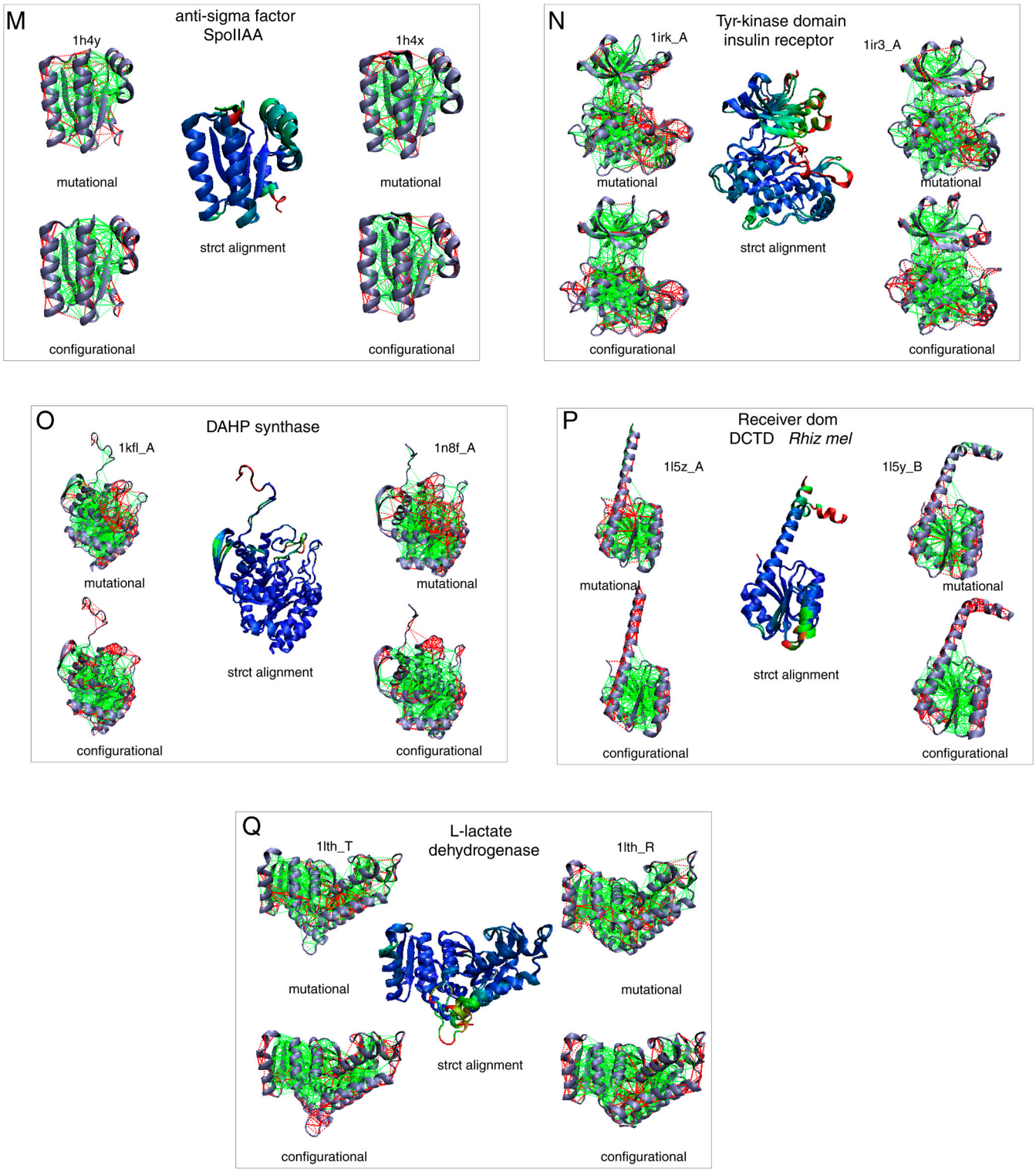


Fig. S1. Gallery of local frustration in allosteric proteins. A structural alignment of both experimentally determined conformations is shown at the center, colored according to the structural deviation (blue low, red high). The individual conformations are shown at the sides. The protein backbone is displayed as ribbons, the direct interresidue interactions with solid lines, and the water-mediated interactions with dashed lines. Minimally frustrated interactions are shown in green, highly frustrated interactions in red, neutral contacts are not drawn. The frustratograms for both the "configurational" (at bottom) and the "mutational" (at top) definitions of local frustration are shown.

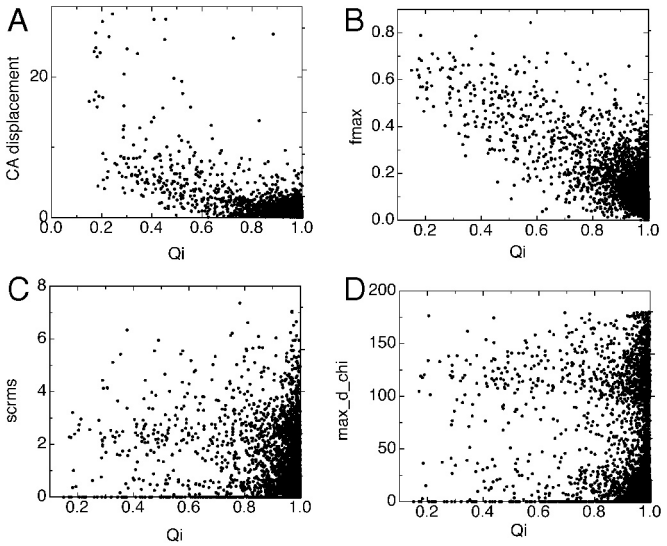


Fig. S2. Comparison of structural deviation metrics. The Q_i value of every residue is compared with other structural metrics developed by Daily and Gray (1), that are based on either backbone (*A* and *B*) or sidechain (*C* and *D*) metrics (1).

1 Daily MD, Gray JJ (2007) Local motions in a benchmark of allosteric proteins. *Proteins* 67:385–399.

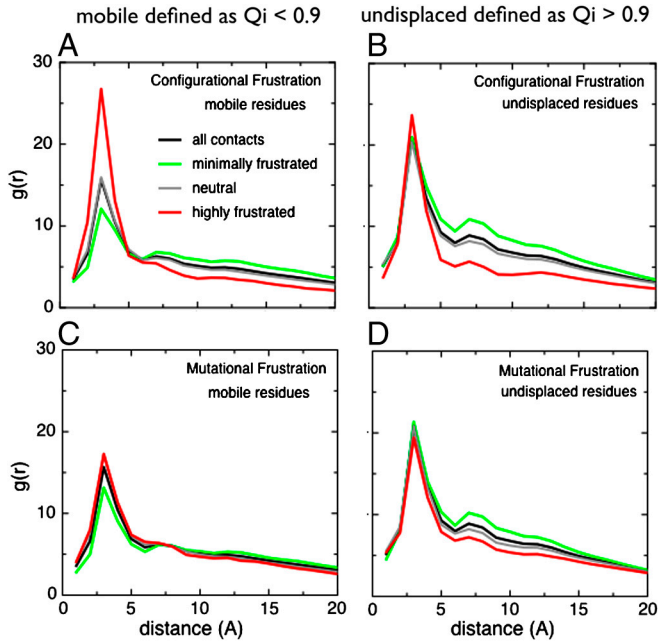


Fig. S3. Local frustration and residue displacement using Q_i. The pair distribution functions between the Cα of the residues classified by displacement and the center of mass of the contacts in different frustration classes was computed. The distributions for all contacts (black), minimally frustrated (green), neutral (gray), or highly frustrated contacts (red), are shown for the mobile (*A* and *C*) or undisplaced (*B* and *D*) residues, using the configurational (*A* and *B*) or mutational (*C* and *D*) frustration indices.

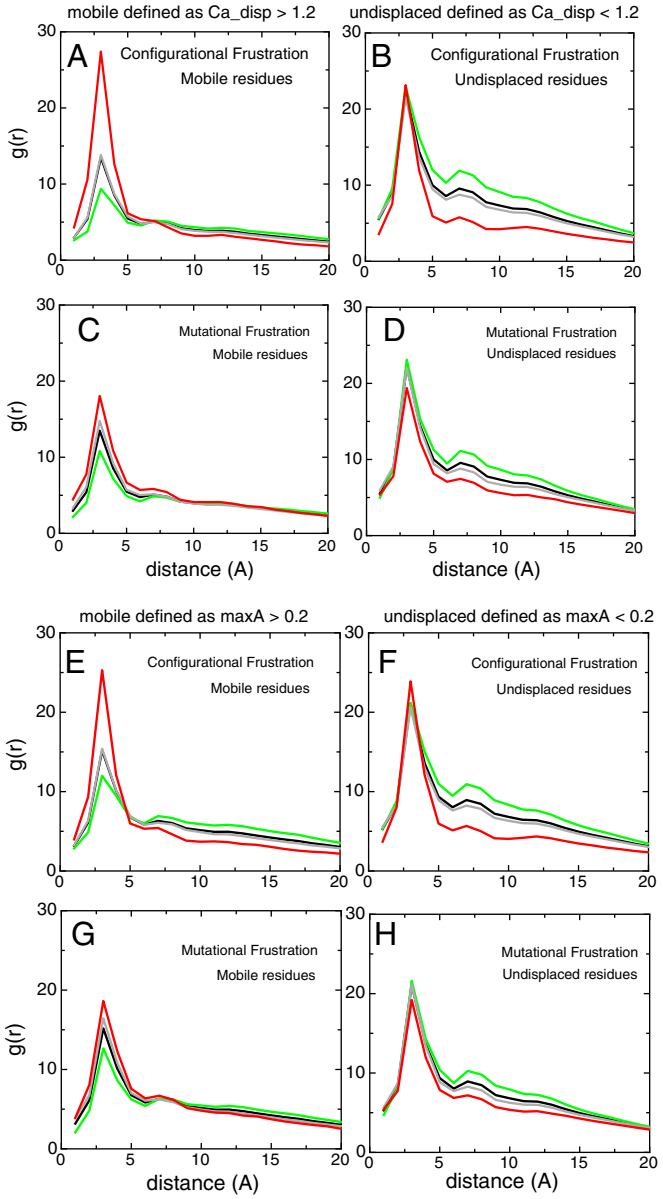


Fig. S4. Local frustration and residue displacement using other metrics. The pair distribution functions between the Cα of the residues classified by displacement using Daily and Gray's definitions of mobile (1) and the center of mass of the contacts in different frustration classes was computed. The distributions for all contacts (black), minimally frustrated (green), neutral (gray), or highly frustrated contacts (red), are shown for the mobile (*A* and *C*) or undisplaced (*B* and *D*) residues, using the configurational (*A* and *B*) or mutational (*C* and *D*) frustration indices. The analogous plots are shown for the definition of mobile based on Fmax (1) in *E*, *F*, *J*, and *K*.

1 Daily MD, Gray JJ (2007) Local motions in a benchmark of allosteric proteins. *Proteins* 67:385–399.

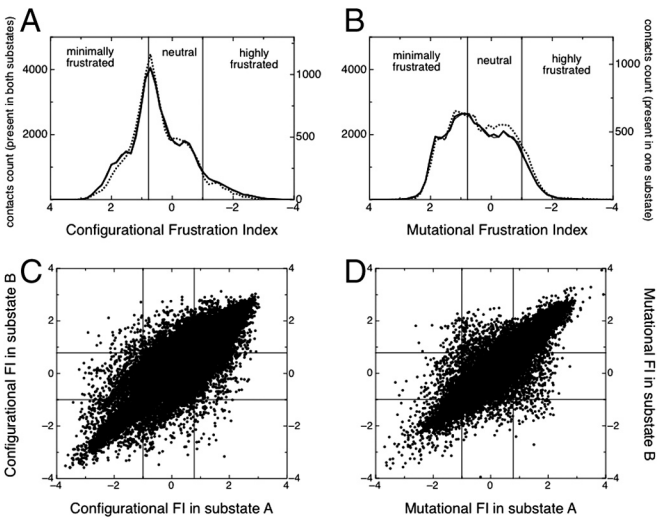


Fig. S5. Local frustration in allosteric pairs. The distribution of the configurational (*A*) and mutational (*B*) frustration indices were calculated for the contacts conserved between substates (solid) or exclusive to one substate (dashed). The vertical lines indicate the cutoff used to define minimally frustrated, neutral, or highly frustrated interactions. For the contacts conserved between substates, the frustration index both in form A or form B are shown as scatter plots (*C* and *D*).

Table S1. Pairs of structures in the database

Open		Closed	
PDB A	Chain	PDB B	Chain
1an0	A	1nf3	A
1cd5	A	1hot	A
1cmb	A	1cma	A
1dbq	A	1wet	A
1e0s	A	2j5x	A
1eyj	A	1eyi	A
1ftn	A	1a2b	A
1g16	A	1g17	A
1hh4	A	1mh1	A
1kao	A	2rap	A
1kfl	A	1n8f	A
1lth	A	1lth	A
1nh8	A	1nh7	A
1oiv	A	1oiw	A
1qr6	A	1pj2	A
1t48	A	1pty	A
1tui	A	1eft	A
1xtq	A	1xts	A
1xxc	A	1xxa	A
2csm	A	1csm	A
2trt	A	1qpi	A
3chy	A	1fqw	A
6pfk	A	4pfk	A

PDB, Protein Data Bank.