Supporting Information

for

Protein Domain-Swapping can be a Consequence of Functional Residues

Nahren Manuel Mascarenhas[1,*] and Shachi Gosavi[1,*]

[1]Simons Centre for the Study of Living Machines

National Centre for Biological Sciences

Tata Institute of Fundamental Research

Bangalore 560065, India

*Correspondence to:

Shachi Gosavi (shachi@ncbs.res.in) or

Nahren Manuel Mascarenhas (mailnahren@gmail.com)

Phone: +91-80-23666105

Fax: +91-80-23636662

Table S1. A summary of simulations performed in this study.		
Simulation	**No. of transitions**	**T_f (reduced units)**
Monellin	102	114.185
stfB	101	115.9
stfB(V48-10x)	98	111.525
stfB(V48-5x)	49	112.9
stfB(V48-1x)	112	115.16
stfB(β1-β2)	82	118.2

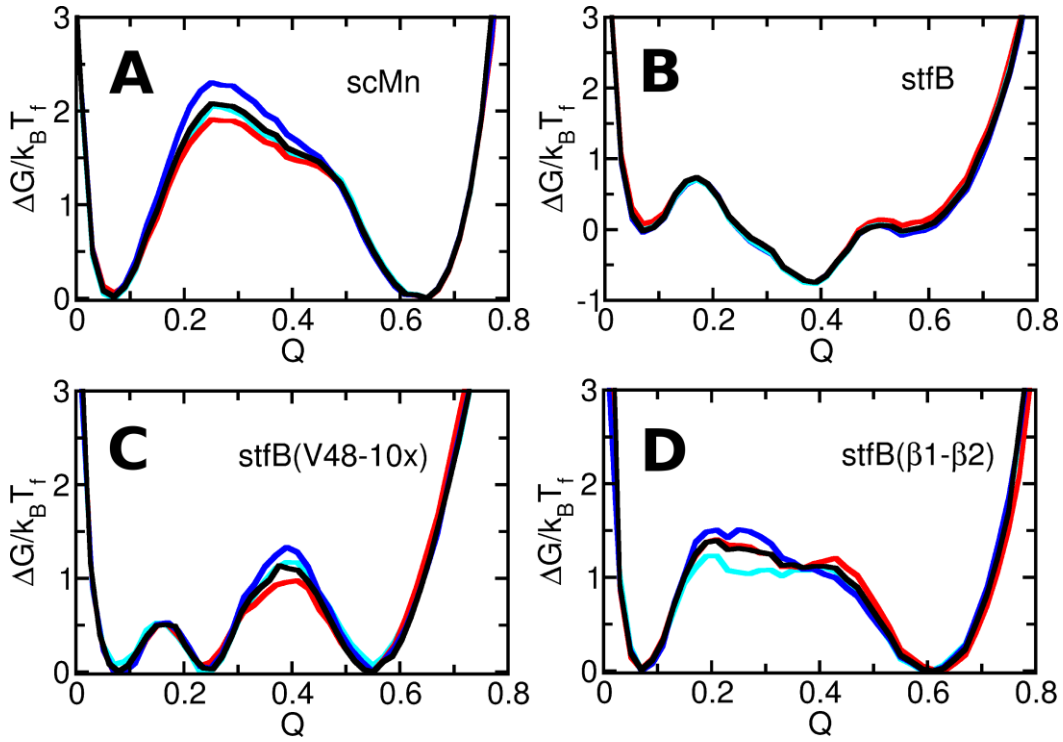

Figure S1. Free energy profiles of segments of the trajectories of scMn (A), stfB (B), stfB(V48-10x) (C), and stfB(β1-β2) (D). The complete folding simulation trajectory for each protein was divided into three approximately equal parts and free energy profiles constructed from each of these segments are shown in cyan, blue and red. Such FEPs (of segments of the trajectory) were used to calculate the error bars shown in Figs. 2A, 3B, 4C and S4. The FEP calculated from the entire dataset of each protein is shown in black and reproduced from Figs. 2A, 3B and 4C.

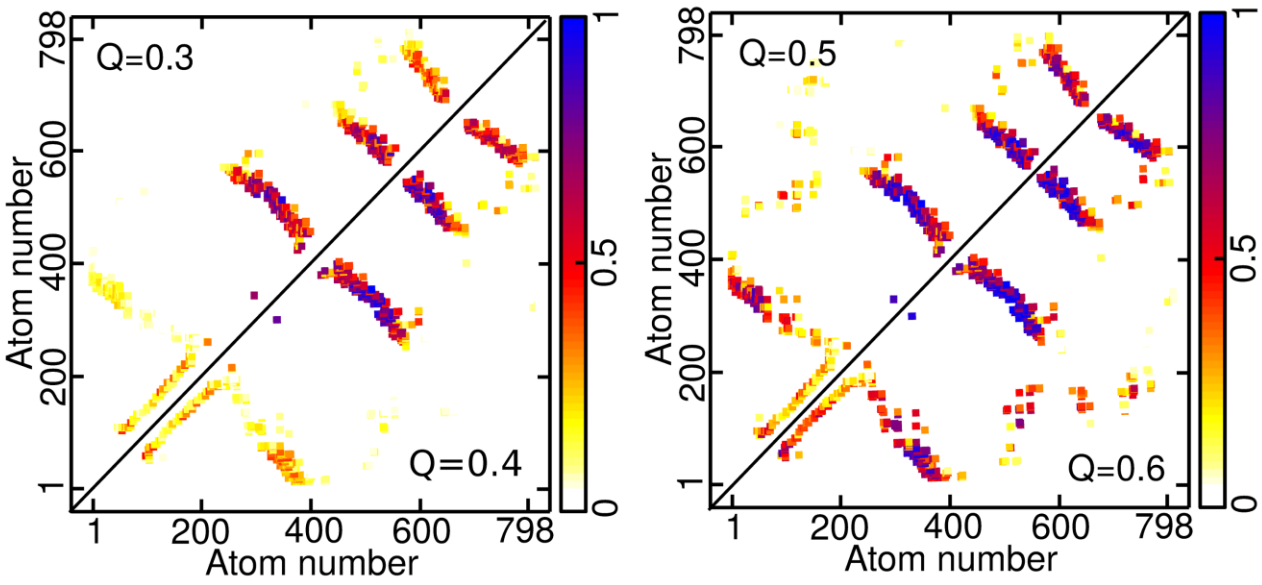

Figure S2. Average contact maps of monellin (scMn) at different values of Q. The colors depict the probability of contact formation and the color scale is given on the right.

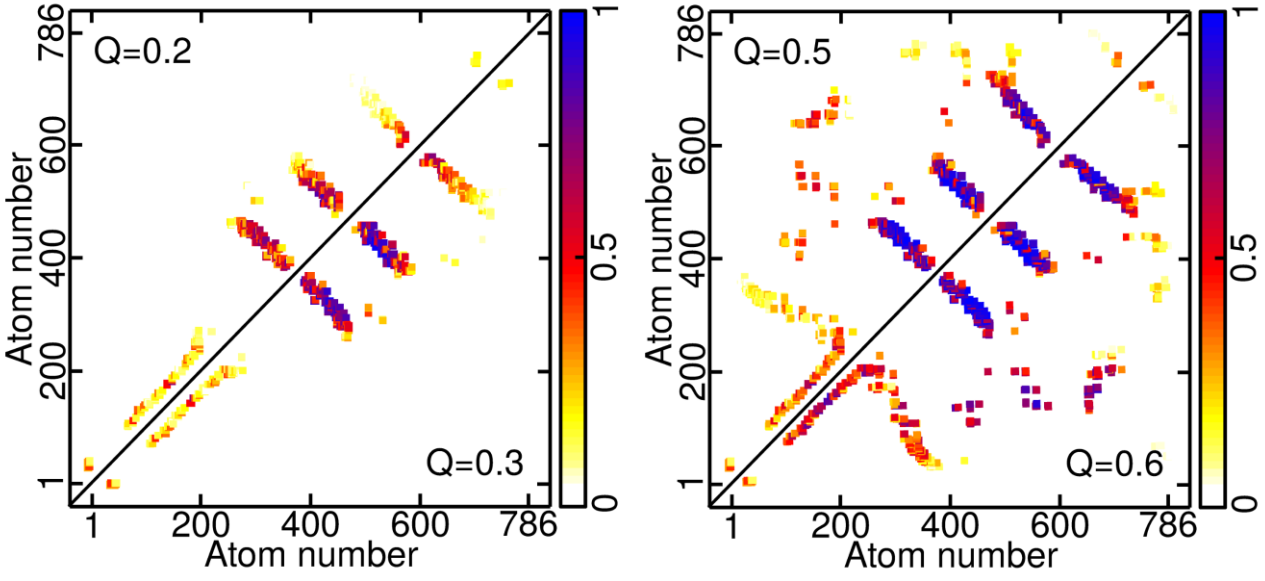

Figure S3. Average contact maps of stfB at different values of Q. The colors depict the probability of contact formation and the color scale is given on the right.

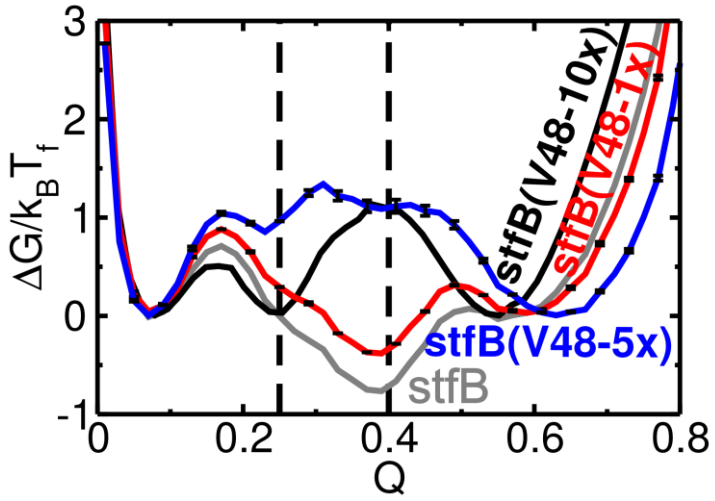

Figure S4. The scaled folding free energies plotted as a function of Q for stfB(V48-1x) and stfB(V48-5x) are compared with those of stfB and stfB(V48-10x). The dashed line on the left marks the position at which the contact map in Fig. 3C (top left) is calculated. The dashed line on the right marks the position at which the contact maps in Fig. 3C (bottom right) and Fig. S5 (second panel, top left) are calculated. There is a small population of both intermediates (marked by the dashed lines) in stfB(V48-5x). The error bars are marked in black on the stfB(V48-1x) and stfB(V48-5x) free energy profiles. The stfB and stfB(V48-10x) profiles are reproduced from Fig. 3B.

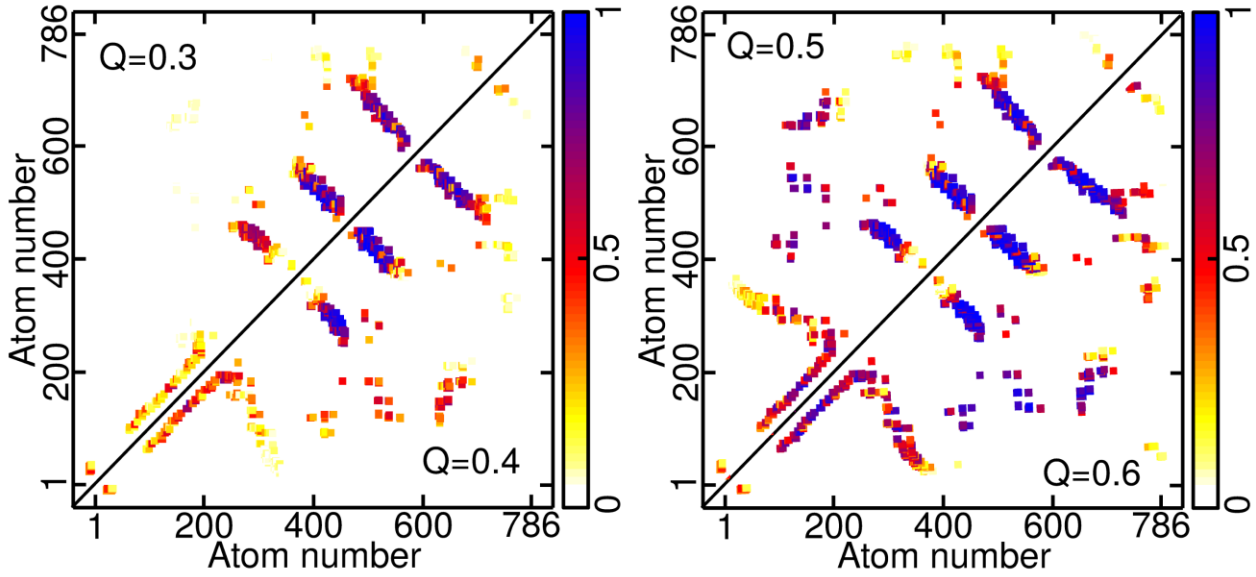

Figure S5. Average contact maps of stfB(V48-10x) at different values of Q. The colors depict the probability of contact formation and the color scale is given on the right.

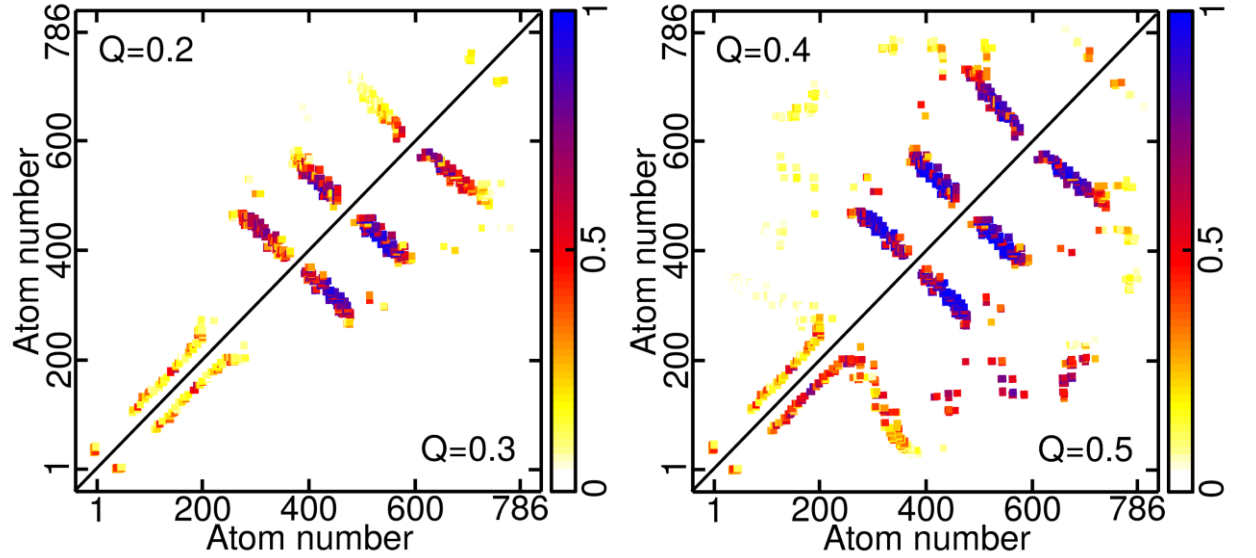

Figure S6. Average contact maps of stfB(V48-1x) at different values of Q. The colors depict the probability of contact formation and the color scale is given on the right.

Table S2. A summary of the results of the stfB symmetrized-SBM simulations.

loop1-hinge	loop1-hinge (one)	two monomers	two monomers (one)	loopA-hinge	loopA-hinge (one)	loop2-hinge	loop3-hinge	uncategorized
8	4	56	7	10	2	5	4	4

Table S3. A summary of the results of the stfB(V48-10x) symmetrized-SBM simulations.

loop1-hinge	loop1-hinge (one)	two monomers	two monomers (one)	loopA-hinge	loopA-hinge (one)	loop2-hinge	loop3-hinge	uncategorized
58	9	2	11	1	1	0	0	18

Table S4. A summary of the results of the stfB(V48-5x) symmetrized-SBM simulations.

loop1-hinge	loop1-hinge (one)	two monomers	two monomers (one)	loopA-hinge	loopA-hinge (one)	loop2-hinge	loop3-hinge	uncategorized
48	10	5	18	4	4	0	0	11

Table S5. A summary of the results of the stfB(V48-1x) symmetrized-SBM simulations.

loop1-hinge	loop1-hinge (one)	two monomers	two monomers (one)	loopA-hinge	loopA-hinge (one)	loop2-hinge	loop3-hinge	uncategorized
18	2	49	5	9	3	5	6	3

A total of 100 quenching runs were performed for each protein and the final snapshots obtained were categorized by visual inspection.

The hinge is that protein segment which links the two swapped domains of a single protein chain.

loop1 is the loop between β2-β3.

loop1-hinge (one): loop1 is the hinge, but only one subunit of the domain-swapped dimer is formed (i.e. β1α1β2β3'β4'β5' is formed).

loopA is the loop between α1-β2

loopA-hinge (one): loopA forms the hinge, but only one subunit of the domain-swapped dimer is formed (i.e β1α1β2'β3'β4'β5' is formed).

two monomers (one): Only one of the monomers is folded.

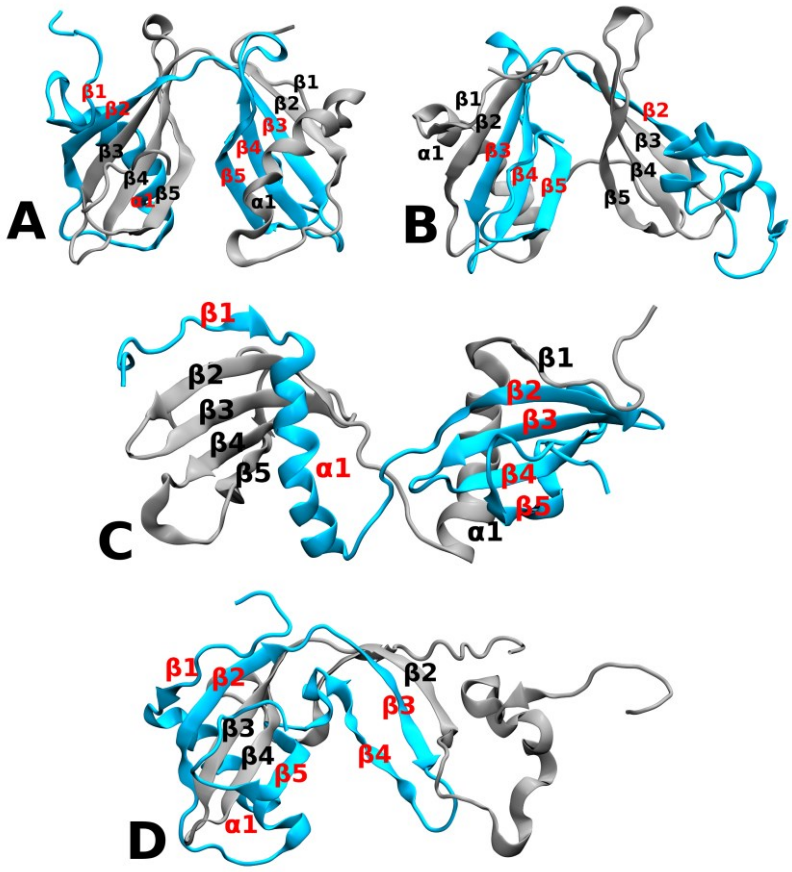

Figure S7. Representative final snapshots from the stfB(V48-10x) symmetrized-SBM simulations are shown to illustrate the different structures obtained. The two chains in the dimer are colored in grey and cyan. (A) Domain swapped dimer where loop1 forms the hinge, similar to the crystal structure of domain-swapped dimer (PDB ID: 2OCT). This structure is often seen in the final snapshots (see Table S3). (B) Only one subunit of the domain swapped dimer (with loop1 as hinge) is formed. In the other subunit α1 and β1 are yet to be formed. Extending such trajectories may allow the formation of the complete domain-swapped dimer (as seen in A). (C) Domain swapped dimer where loopA forms the hinge. (D) A few of the final snapshots cannot be easily structurally categorized and these have been labelled uncategorized in Table S3. Here, we show an example of such a structure where only β3 and β4 are exchanged.

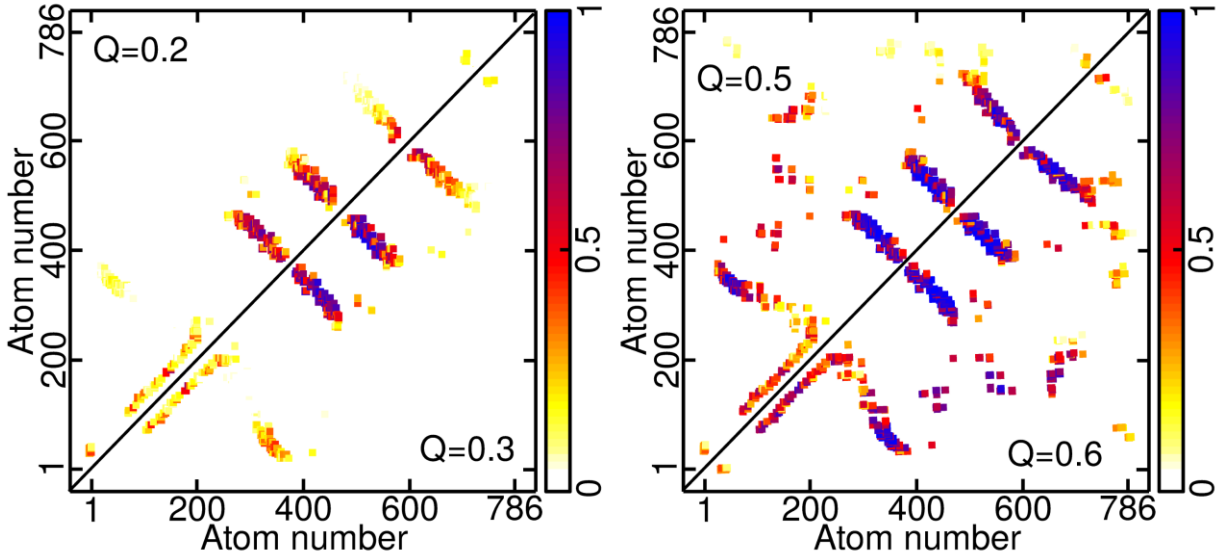

Figure S8. Average contact maps of stfB(β1-β2) at different values of Q. The colors depict the probability of contact formation and the color scale is given on the right.

Table S6. Residues of stefin-B which are within 4.5Å of papain in the papain-stefin-B complex (PDB ID: 1STF) are marked in Fig. 4E. The residue IDs here are renumbered according to those present in the monomer structure of stefin-B (PDB ID: 4N6V chain 0). However, residues in the N^{ter} region and loop1 contribute more to the binding interface because only these residues have three or more contacts with papain.

Residues at the N^{ter} region	1, 2, 3, 4, 5, 6
Residues part of loop1	46, 47, 48, 49, 51, 52
Residues prior to loop3	68, 70
Residues part of loop3	73, 74, 75
Residues at the C^{ter} region	97

Functional residues in monellin (residues that are marked on Fig. 4F): 6, 7, 9, 13, 16, 36, 39, 43, 72, 88, 92-96.

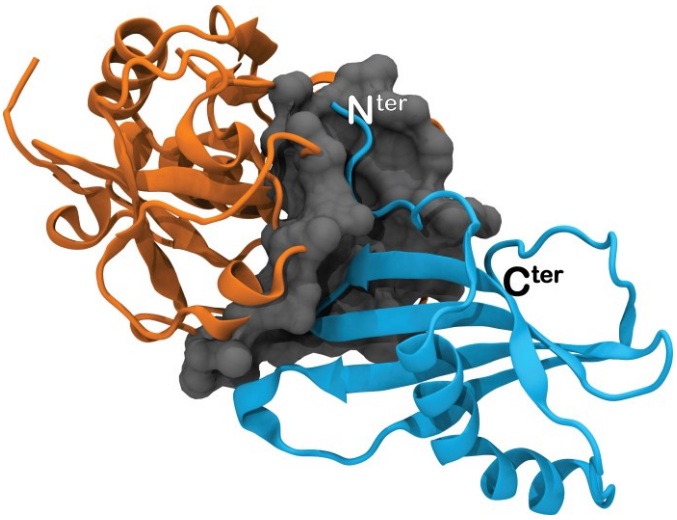

Figure S9. The crystal structure of the papain-stefin-B complex is shown here to highlight residues of papain (orange) and stefin-B (cyan) which are involved in binding. The surface shown (grey) corresponds to those residues of papain which are within 4.5 Å of stefin-B. It is apparent that residues from the N-terminus (N^{ter}) and loop1 of stefin-B interact extensively with papain.