

# Supplementary Information

## Is Post-Translational Folding More Efficient Than Refolding from a Denatured State: A Computational Study

Quyen V. Vu<sup>1</sup>, Daniel A. Nissley<sup>2</sup>, Yang Jiang<sup>3</sup>, Edward P. O'Brien<sup>3, 4, 5\*</sup> and Mai Suan Li<sup>1, 6\*</sup>

<sup>1</sup>Institute of Physics, Polish Academy of Sciences; Al. Lotnikow 32/46, 02-668 Warsaw, Poland.

<sup>2</sup>Department of Statistics, University of Oxford, Oxford, OX1 3LB, UK.

<sup>3</sup>Department of Chemistry, Pennsylvania State University, University Park, PA 16802, USA.

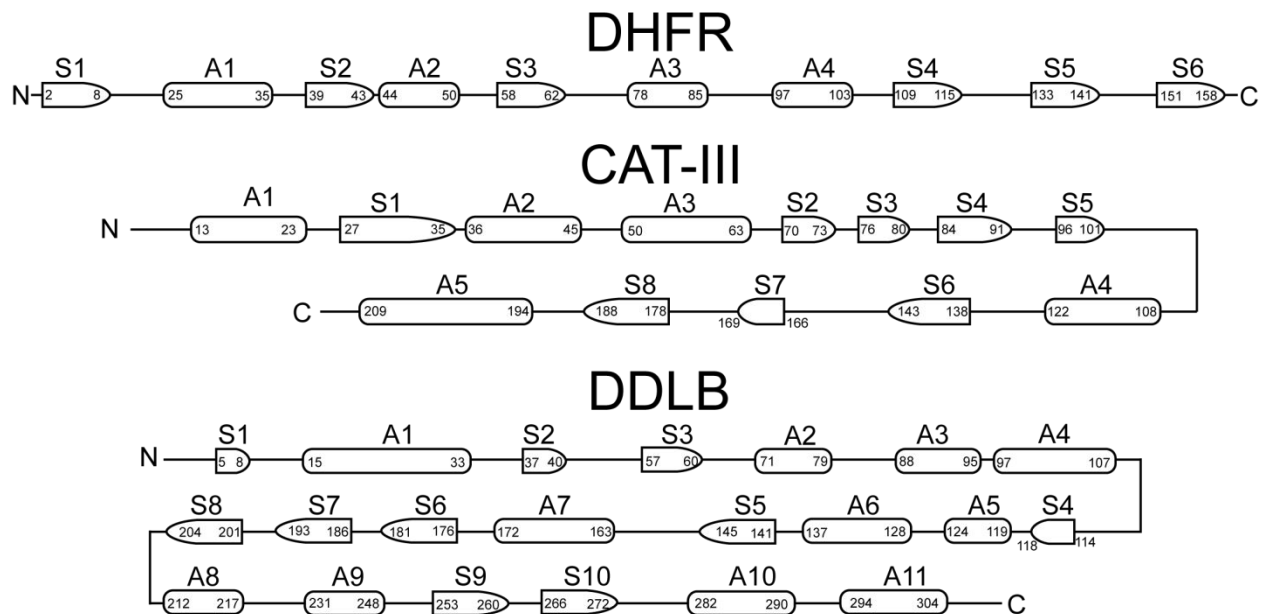
<sup>4</sup>Bioinformatics and Genomics Graduate Program, The Huck Institutes of the Life Sciences, Pennsylvania State University, University Park, PA 16802, USA.

<sup>5</sup>Institute for Computational and Data Sciences, Pennsylvania State University, University Park, PA 16802, USA.

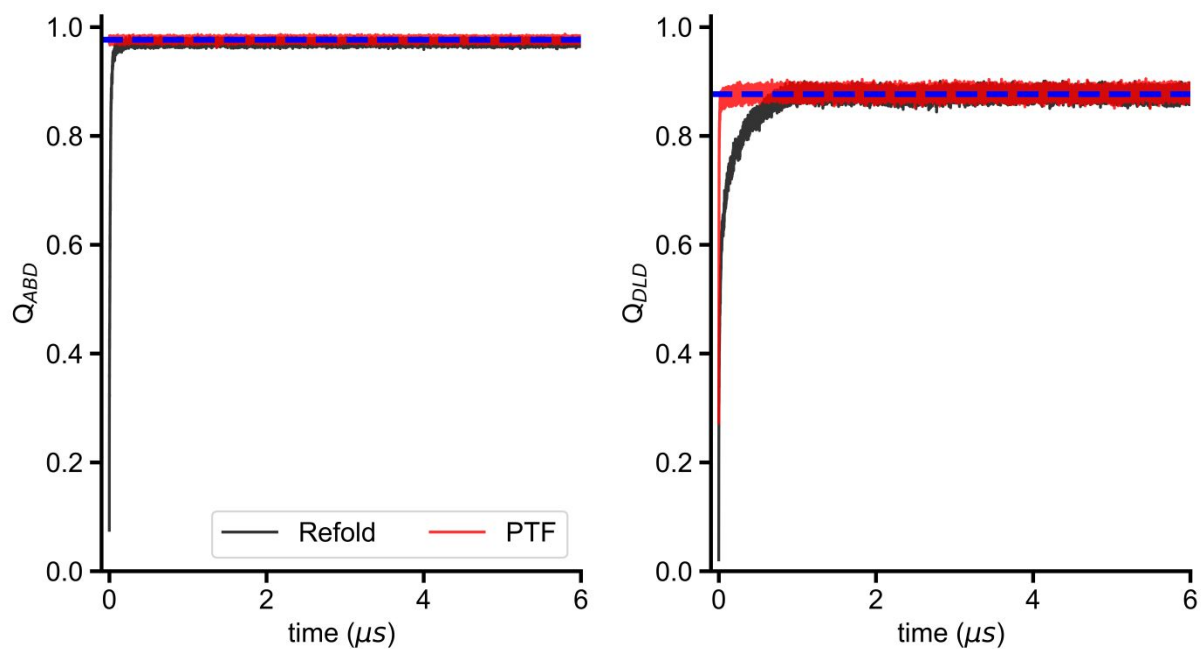
<sup>6</sup>Institute for Computational Sciences and Technology; Quang Trung Software City, Tan Chanh Hiep Ward, District 12, Ho Chi Minh City, Vietnam.

\* To whom correspondence should be addressed: [epo2@psu.edu](mailto:epo2@psu.edu), [masli@ifpan.edu.pl](mailto:masli@ifpan.edu.pl)

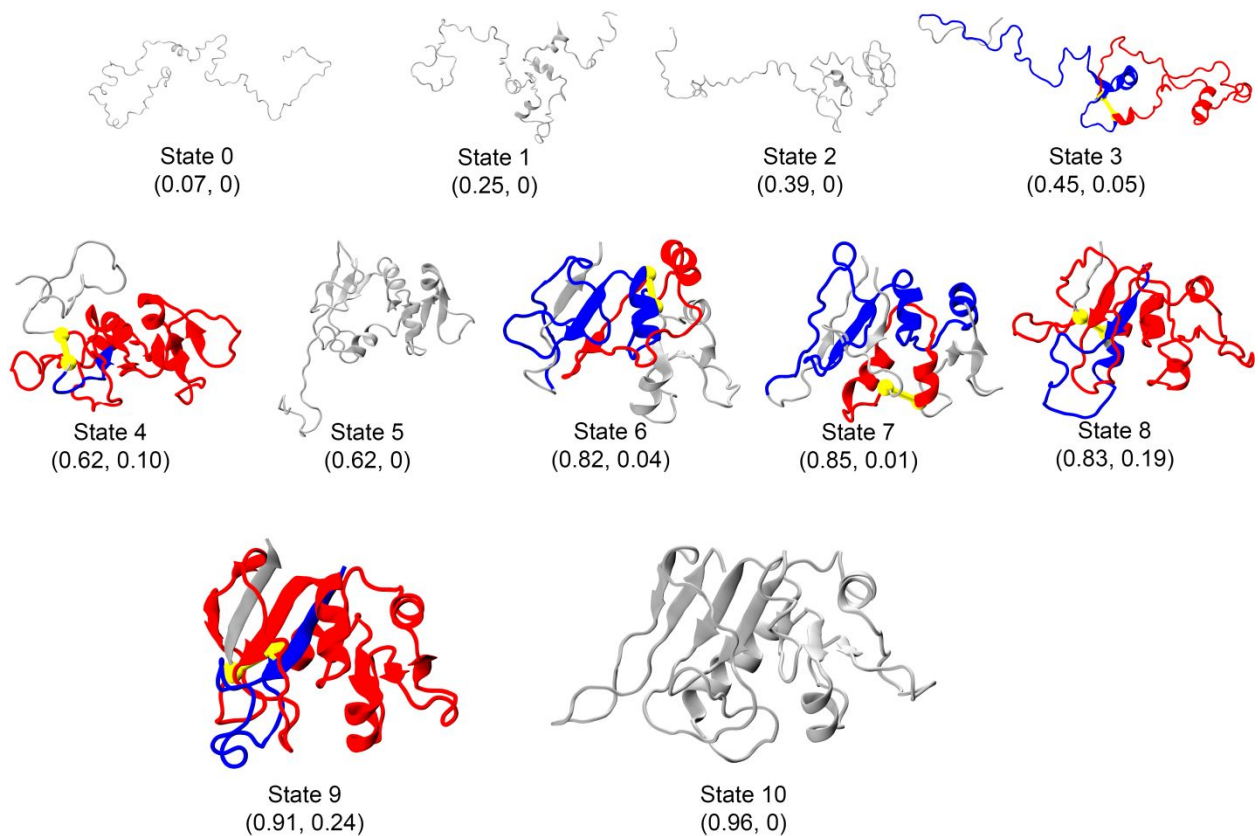
## Supplementary Figures



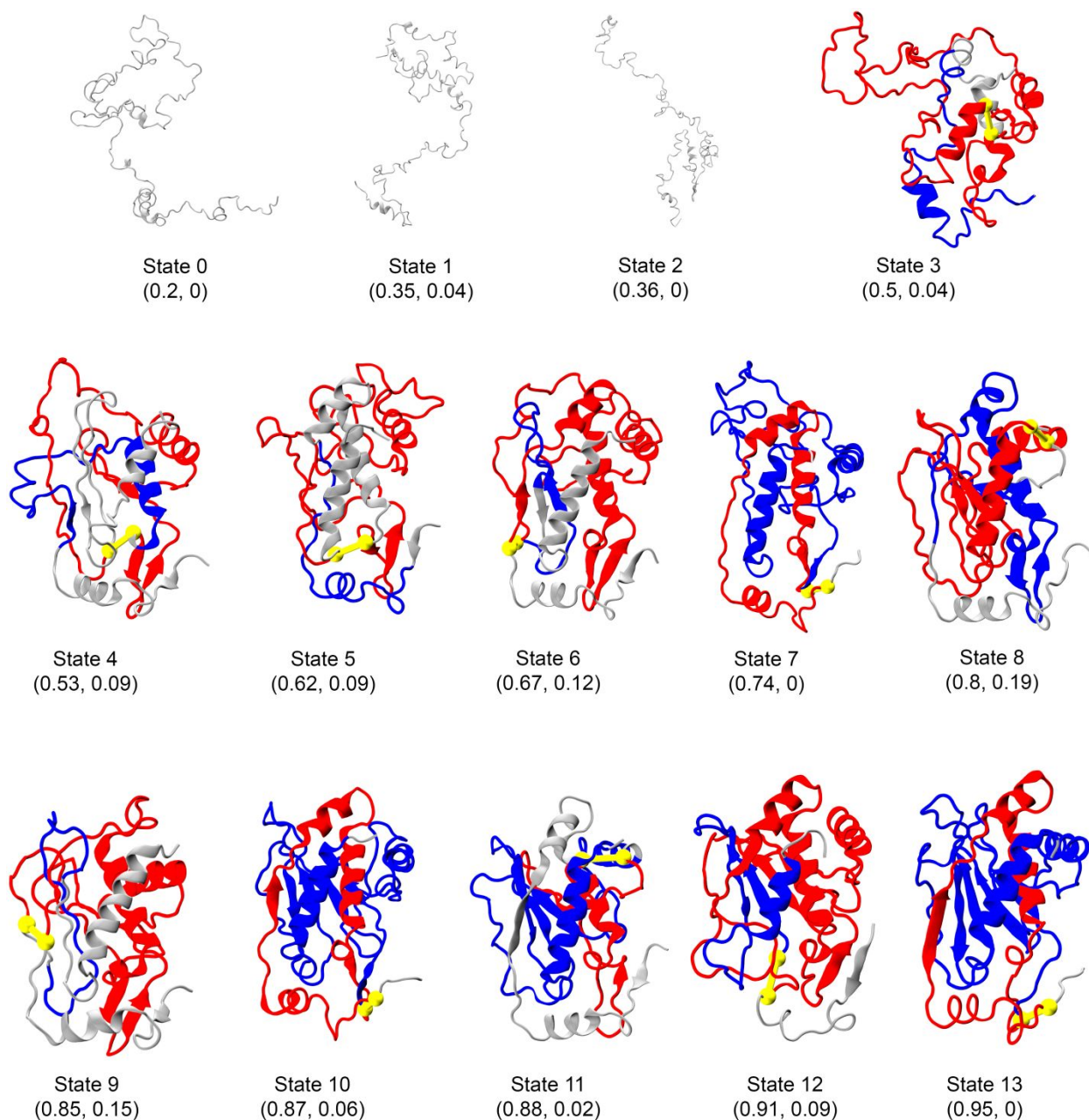
**Figure S1: Secondary structure elements in DHFR, CAT-III, and DDLB.**  $\alpha$ -helices and  $\beta$ -sheets are labeled as A and S. Noted that the secondary structures are determined by Stride and we only show segments with more than 4 residues.



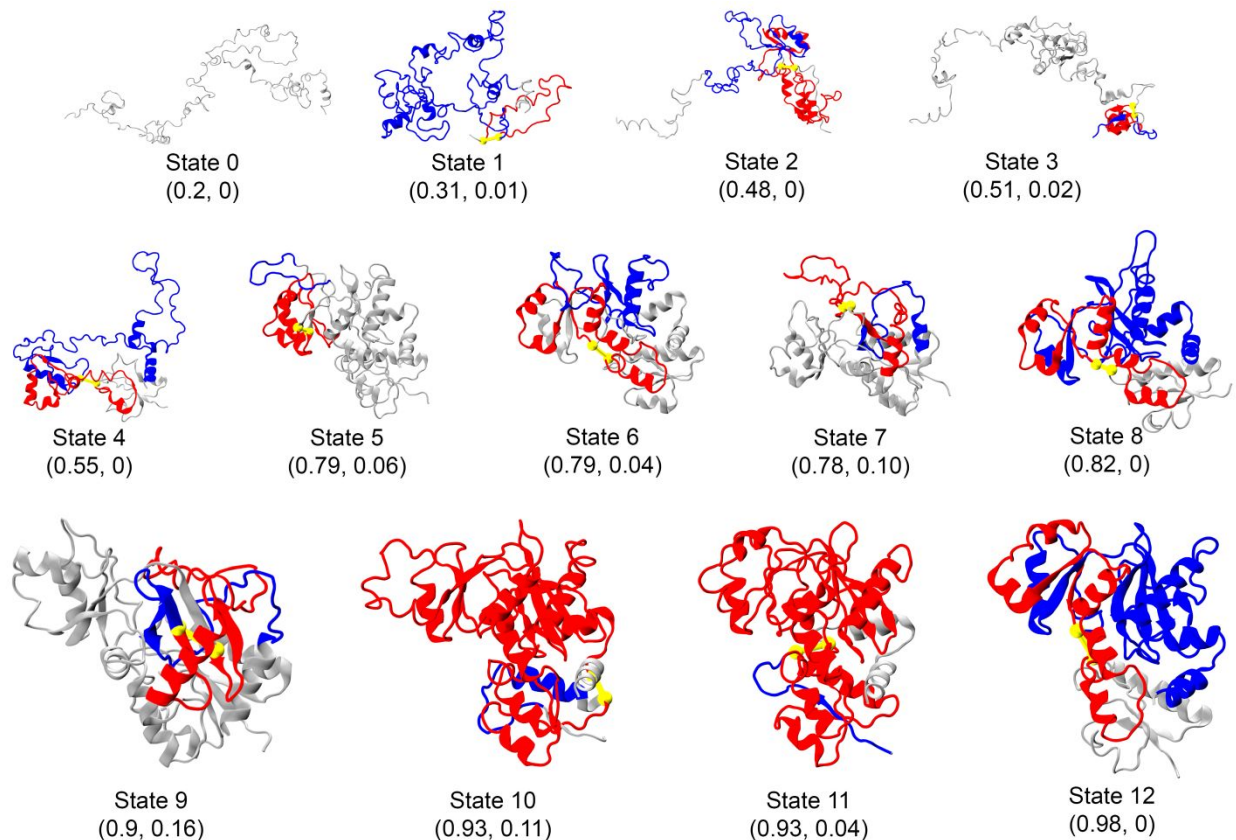
**Figure S2: Fraction of native contacts versus time computed from folded trajectories of DHFR protein (un-normalized).** (Left): ABD domain, (right) DLD domain. The blue-dashed line presents the threshold for each domain determined from the native simulation.



**Figure S3: Representative structures of the metastable states from clustering of DHFR.** The all-atom backmapped structures are presented with  $(Q, G)$  values. Entanglement in structure (if exists) is highlighted with the loop (red color) and is closed by the native contact (yellow sphere and yellow bond), and the threading segment is colored in blue. The metastable state indices and the  $(Q, G)$  values of the most probable microstates in that metastable state are shown below each structure.



**Figure S4: Representative structures of the metastable states from clustering of CAT-III.** The all-atom backmapped structures are presented with  $(Q, G)$  values. Entanglement in structure (if exist) is highlighted with the loop (red color) and is closed by the native contact (yellow sphere and yellow bond), and the threading segment is colored in blue. The metastable state indices and the  $(Q, G)$  values of the most probable microstates in that metastable state are shown below each structure.



**Figure S5: Representative structures of the metastable states from clustering of DDLB.** The all-atom backmapped structures are presented with  $(Q, G)$  values. Entanglement in structure (if exist) is highlighted with the loop (red color) and is closed by the native contact (yellow sphere and yellow bond), and the threading segment is colored in blue. The metastable state indices and the  $(Q, G)$  values of the most probable microstates in that metastable state are shown below each structure.

## Supplementary Tables

**Table S1: Wild-type mRNA template for three enzymes in this study.**

Protein	mRNA template - Wildtype translation
DHFR	AUGAUCAGUCUGAUUJGCGGCGUJAGCGGUAGAUCGCGUUAUCGGCAUGGAAAACG CCAUGCCGUGGAACCUGCCUGCCGAUCUCGCCUGGUUUAAACGCAACACCUUAAA UAAACCCGUGAUUAUGGGCCGCCAUACCUGGGAAUCAAUCCGUCGUCGGUUGCCA GGACGCAAAAUAUUAUCCUCAGCAGUCAACCGGGUACGGACGAUCGCGUAACGU GGGUGAAGUCGGUGGAUGAAGCCAUCGCGGCGUGUGGUGACGUACCAGAAAUCAU GGUGAUUGGCGGCGGUCGCGUUUAUGAACAGUUCUUGCCAAAAGCGCAAAAACUG UAUCUGACGCAUAUCGACGCAGAAGUGGAAGGCGACACCCAUUUCCCGGAUUACG AGCCGGAUGACUGGGAUUCGGUAUUCAGCGAAUCCACGAUGCUGAUGCAGAA CUCUCACAGCUAUUGCUUUGAGAUUCUGGAGCGGCGGUA

<p>CAT-III</p>	<p>AUGAACUAUACAAAAUUUGAUGUAAAAAAUUGGGUUCGCCGUGAGCAUUUUGAGUU  UUUUCGGCAUCGUUUACCAUGUGGUUUUAGCUUAACAAGCAAAAUUGAUUUCACGA  CGUUAAAAAAGUCAUUGGAUGAUUCAGCGUAUAAGUUUUUUCGGUAAUGAUCUUAU  CUGAUUGCUCAGGCCGUGAAUCAUUUGAUGAGUUUGAGAAUGGCGAUAAAAGAUG  AUGAAUUGAUCGUAUGGGAUUCAGUCGACCCACAUAUCACCGUAUCCAUCUAAAGAA  ACAGAGACAUUUUCAGCACUGAGUUGCCCAUACUCAUCCGAUUAUUGAUCAAUUUAU  GGUGAAUUAUUUAUCGGUAAUGGAACGUUAUAAAAGUGAUACCAAGUUAUUUCCUC  AAGGGGUAACACCAGAAAUCAUUUAAAUAUUUCAGCAUUACCUUGGGUUAUUUUU  GAUAGCUUUAAUUUAAAUGUUGCUAAUUUUACCGAUUAUUUUGCACCCAUUAUAAC  AAUGGCAAAAUAUCAGCAAGAAGGGGGAUAGACUGUUUAUUGCCGCUCUCAGUACAGG  UUCAUCAUGCAGUUUGUGAUGGCUUCCAUGUUGCACGCUUUUAUUAUCGGCUACA  AGAGUUGUGUAACAGUAAAUAUUAAAUA</p>
<p>DDL B</p>	<p>AUGACUGAUAAAAUCGCCGUCCUGUUGGGUGGGACCUCGCGUGAGCGGGAAGUUU  CUCUGAAUUCUGGCGCAGCGGUGUAGCCGGACUGCGUGAAGGCGGUUUUGACG  CGUAUCCUGUCGACCCGAAAGAAGUCGACGUGACGCAACUGAAGUCGAUUGGCUU  UCAGAAAGUGUUUAUCGCGCUACACGGUCGCGGCGGUGAAGAUGGUACGCUGCAG  GGGAUGCUCGAGCUGAUGGGCUUGCCUUAUACCGGAAGCGGAGUGAUGGCAUCU  GCGCUUUCAAUGGAUAAACUACGCAGCAAACUUCUAUGGCAAGGUGCCGGUUUAC  CGGUCGCGCCGUGGGUAGCGUUAACCCGCGCAGAGUUUGAAAAAGGCCUGAGCGA  UAAGCAGUJAGCAGAAUUUCUGCUCUGGGUUUGCCGGUUAUCGUUAAGCCGAGC  CGCGAAGGUUCCAGUGUGGGAAUGUCAAAAGUAGUAGCAGAAAUGCUCUACAAGA  UGCAUUAAGAUUGGCAUUUCAGCACGAUGAAGAAGUAUUGAUUGAAAAAUGGCUAA  GUGGGCCGGAGUUCACGGUUGCGAUACUCGGUGAAGAAAUUUUACCGUCAAUACG  UAUUCACCGUCCGGAACCUUCUAUGAUUAUGAGGCGAAGUAUCUCUCUGAUGAGA  CACAGUAUUUCUGCCCCGAGGUCUGGAAGCGUCACAAGAGGCCAAUUUGCAGGC  AUUAGUGCUGAAAGCAUGGACGACGUUAGGUUGCAAAGGAUGGGGACGUUUUGAC  GUUAUGCUGGACAGCGAUGGACAGUUUUUAUCUGCUGGAAGCCAUAUACCUACCCGG  GUAUGACCAGCCACAGCCUGGUGCCGAUGGCGGCACGUCAGGCAGGUUAUGAGCUU  CUCGCAGUUGGUAGUACGAAUUCUGGAACUGGCGGACUAA</p>

**Table S2: Structural definitions as the pair of secondary structures (shown in Figure S1) in the native state of proteins.**

Protein	Pair		#native contacts	Pair		#native contacts	Pair		#native contacts
DHFR	<b>C1</b>	S <sub>5</sub> -S <sub>6</sub>	30	<b>C6</b>	S <sub>1</sub> -A <sub>4</sub>	3	<b>C11</b>	A <sub>1</sub> -S <sub>2</sub>	2
	<b>C2</b>	S <sub>4</sub> -S <sub>6</sub>	25	<b>C7</b>	S <sub>2</sub> -A <sub>2</sub>	3	<b>C12</b>	S <sub>2</sub> -A <sub>3</sub>	2
	<b>C3</b>	S <sub>1</sub> -S <sub>4</sub>	21	<b>C8</b>	S <sub>2</sub> -A <sub>4</sub>	3	<b>C13</b>	A <sub>3</sub> -A <sub>4</sub>	2
	<b>C4</b>	S <sub>2</sub> -S <sub>3</sub>	16	<b>C9</b>	A <sub>2</sub> -S <sub>3</sub>	3			
	<b>C5</b>	S <sub>3</sub> -A <sub>3</sub>	7	<b>C10</b>	S <sub>1</sub> -A <sub>1</sub>	2			
CAT-III	<b>C1</b>	S <sub>1</sub> -S <sub>8</sub>	32	<b>C6</b>	S <sub>6</sub> -S <sub>7</sub>	12	<b>C11</b>	A <sub>3</sub> -S <sub>4</sub>	4
	<b>C2</b>	S <sub>4</sub> -S <sub>5</sub>	22	<b>C7</b>	A <sub>3</sub> -A <sub>5</sub>	9	<b>C12</b>	A <sub>3</sub> -A <sub>4</sub>	4
	<b>C3</b>	S <sub>4</sub> -S <sub>6</sub>	16	<b>C8</b>	A <sub>2</sub> -S <sub>8</sub>	7	<b>C13</b>	A <sub>3</sub> -S <sub>6</sub>	2
	<b>C4</b>	S <sub>7</sub> -S <sub>8</sub>	14	<b>C9</b>	A <sub>2</sub> -A <sub>5</sub>	7	<b>C14</b>	S <sub>5</sub> -A <sub>4</sub>	2
	<b>C5</b>	S <sub>2</sub> -S <sub>3</sub>	13	<b>C10</b>	S <sub>2</sub> -A <sub>5</sub>	4			
DDLB	<b>C1</b>	S <sub>7</sub> -S <sub>9</sub>	26	<b>C11</b>	S <sub>5</sub> -A <sub>7</sub>	6	<b>C21</b>	A <sub>4</sub> -S <sub>10</sub>	3
	<b>C2</b>	S <sub>9</sub> -S <sub>10</sub>	24	<b>C12</b>	A <sub>3</sub> -S <sub>10</sub>	5	<b>C22</b>	S <sub>4</sub> -A <sub>5</sub>	3
	<b>C3</b>	S <sub>4</sub> -S <sub>6</sub>	16	<b>C13</b>	S <sub>4</sub> -A <sub>6</sub>	5	<b>C23</b>	A <sub>6</sub> -S <sub>6</sub>	3
	<b>C4</b>	S <sub>5</sub> -S <sub>6</sub>	16	<b>C14</b>	S <sub>7</sub> -A <sub>10</sub>	5	<b>C24</b>	A <sub>5</sub> -S <sub>6</sub>	2
	<b>C5</b>	A <sub>1</sub> -A <sub>11</sub>	15	<b>C15</b>	A <sub>10</sub> -A <sub>11</sub>	5	<b>C25</b>	S <sub>7</sub> -A <sub>9</sub>	2
	<b>C6</b>	S <sub>7</sub> -S <sub>8</sub>	13	<b>C16</b>	S <sub>1</sub> -A <sub>1</sub>	4	<b>C26</b>	S <sub>8</sub> -A <sub>10</sub>	2
	<b>C7</b>	S <sub>1</sub> -S <sub>2</sub>	12	<b>C17</b>	A <sub>2</sub> -A <sub>3</sub>	4	<b>C27</b>	A <sub>9</sub> -S <sub>9</sub>	2
	<b>C8</b>	S <sub>1</sub> -S <sub>3</sub>	12	<b>C18</b>	A <sub>3</sub> -A <sub>4</sub>	4	<b>C28</b>	A <sub>9</sub> -S <sub>10</sub>	2
	<b>C9</b>	A <sub>1</sub> -S <sub>2</sub>	7	<b>C19</b>	A <sub>7</sub> -S <sub>6</sub>	4			
	<b>C10</b>	A <sub>4</sub> -A <sub>9</sub>	6	<b>C20</b>	A <sub>3</sub> -A <sub>9</sub>	3			