Biophysical Journal

Supporting Material

Dead-End Elimination with a Polarizable Force Field Repacks PCNA Structures

Stephen D. LuCore,[1] Jacob M. Litman,[2] Kyle T. Powers,[2] Shibo Gao,[2] Ava M. Lynn,[1] William T. A. Tollefson,[1] Timothy D. Fenn,[3] M. Todd Washington,[2] and Michael J. Schnieders[1,2,*]

[1]Department of Biomedical Engineering and [2]Department of Biochemistry, University of Iowa, Iowa City, Iowa; and [3]Boehringer Ingelheim, Ridgefield, Connecticut

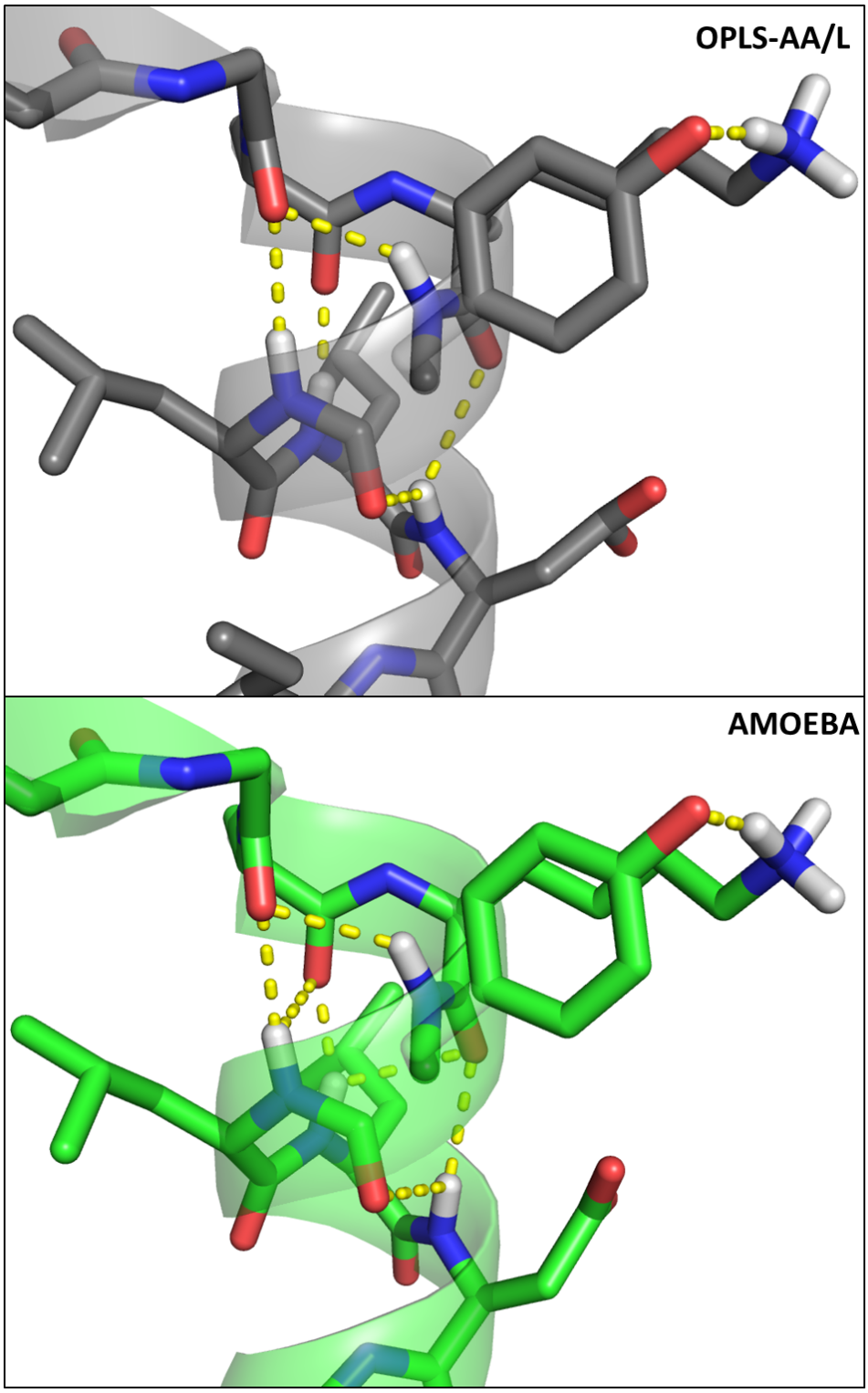
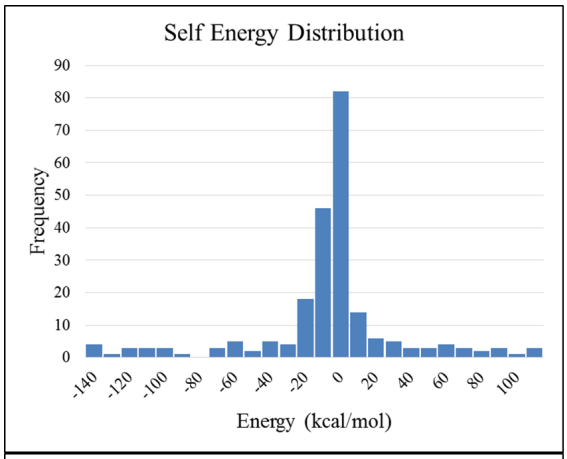
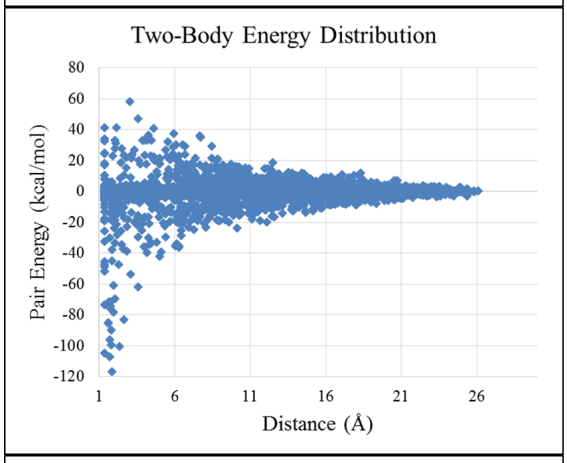
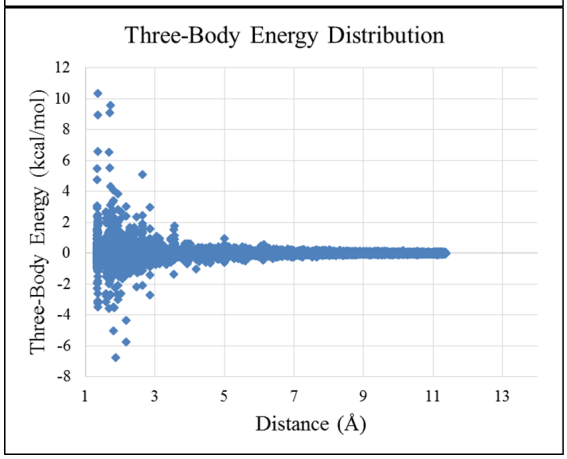
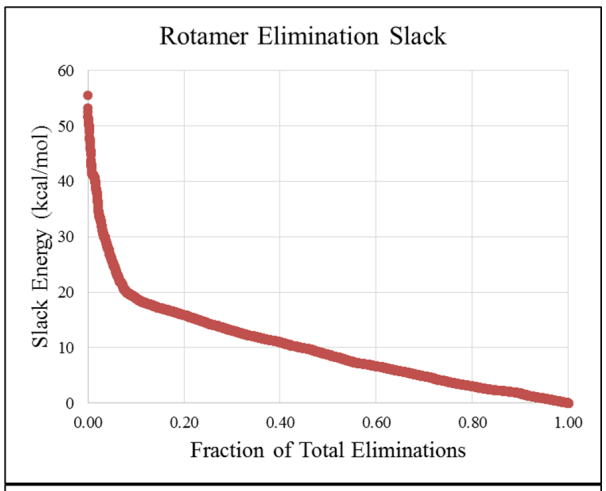


Figure S1. Hydrogen bonding for an α-helix from wild type PCNA is shown after application of OPLS-AA/L pairwise DEE and AMOEBA many-body DEE. Hydrogen atoms not involved in hydrogen bonding have been hidden for clarity. The AMOEBA model (lower panel) shows two additional hydrogen bonds not present in the OPLS-AA/L model (upper panel) due to backbone nitrogen atoms of residue *i* donating to both the residue *i*+3 and *i*+4 carbonyl oxygen atoms.









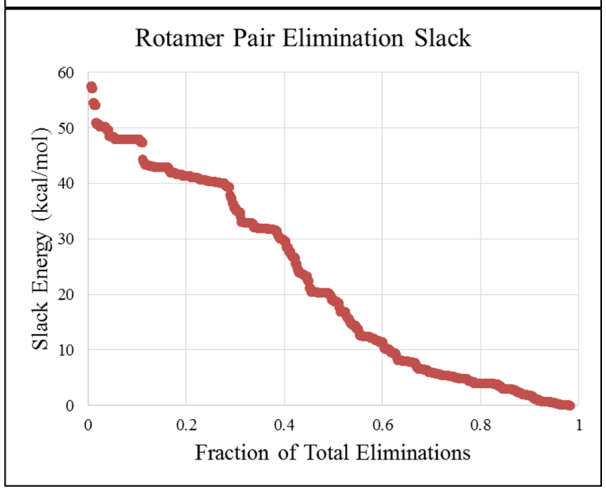


Figure S2. Shown to the left are distributions of self, 2-body, and 3-body energies from the wild type PCNA data as a function of distance. Shown to the right are sorted distributions of slack for rotamer and rotamer pair eliminations under AMOEBA many-body DEE for wild type PCNA (2 values for rotamer slack and 28 values for rotamer pair slack are greater than 60 kcal/mol and were not included). The maximum absolute 3-body energy is 10.4 kcal/mol, which is greater than 57% of the rotamer slack energies and 37% of the rotamer pair slack energies. Although it is infeasible to compute the distribution of all 4-body energies (there are more than 100 million), a subset of 1.9 million establishes their maximum absolute value to be approximately 1 kcal/mol (see Figure S3 below). We note that only 0.06% of rotamer slack energies and 0.06% of rotamer pair slack energies are less than 1 kcal/mol.

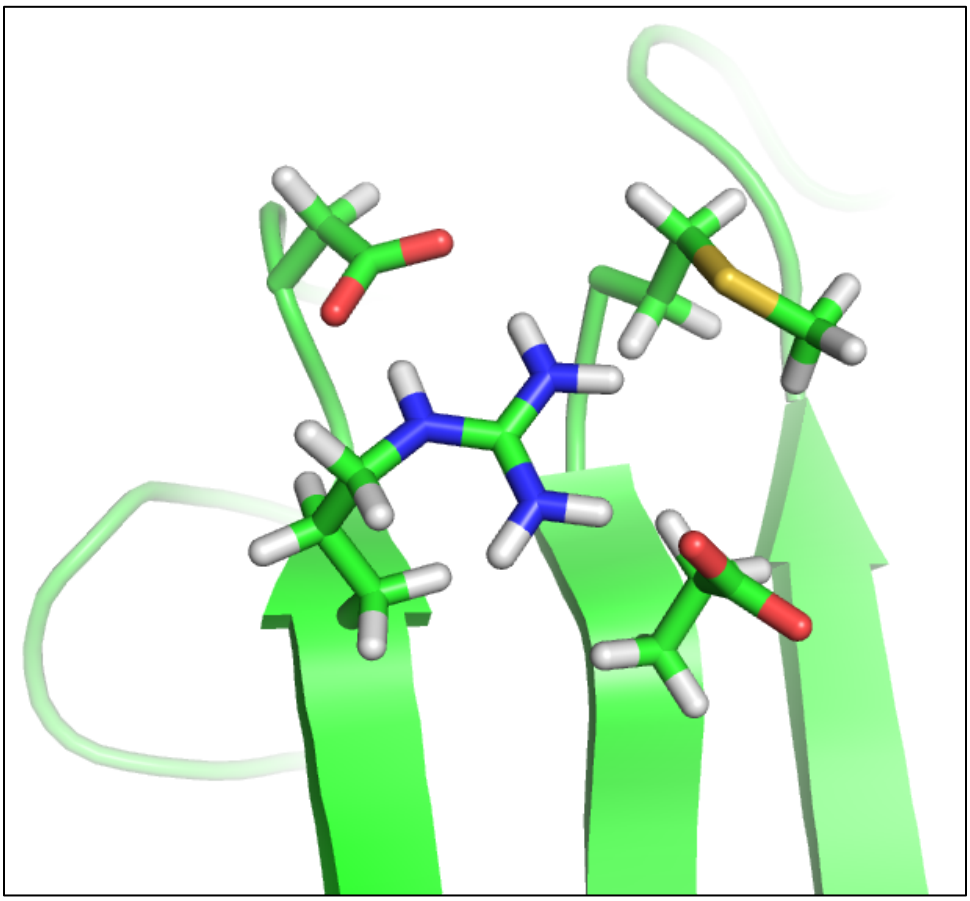


Figure S3. The shown collection of 4 residues from wild type PCNA produced the largest absolute magnitude 4-body energy among the subset of 1.9 million evaluated that contained the N-terminal methionine (-0.98 kcal/mol). The second largest 4-body energy identified was 0.39 kcal/mol and the average 4-body energy was -9e-07 kcal/mol.

Table S1. Shown is quantification of hydrogen bonding at the PCNA subunit interface. Intra-molecular hydrogen bonds are those spanning either βH_1-βI_1 or βC_1-βD_1. Inter-molecular hydrogen bonds are those spanning the subunit interface βI_1-βD_1. The original structural models do not show a clear trend in hydrogen bonding for the E113G and G178S mutants relative to wild type. However, the AMOEBA DEE repacked models reveal an increase in intra-subunit hydrogen bonding in the mutant structures that is consistent with reduced trimer stability.

Data		**Intra-molecular**		**Inter-molecular**		**Delta**
Set	**Model**	**Total**	**Change**	**Total**	**Change**	**Change**
WT	Original	22		5		
	PDB_Redo	20		5		
	OPLS-AA/L	29		5		
	AMOEBA	28		5		
3F1W	Original	22	0	4	-1	+1
(E133G)	PDB_Redo	22	2	4	-1	+3
	OPLS-AA/L	32	3	5	0	+3
	AMOEBA	31	3	6	+1	+2
3GPM	Original	22	0	4	-1	+1
(G178S)	PDB_Redo	19	-1	5	0	-1
	OPLS-AA/L	33	4	6	+1	+3
	AMOEBA	34	6	5	0	+6

Table S2. Shown are coordinate RMSDs relative to deposited coordinates and mean deviations in bonds lengths and angle bends.

Data			RMSD (Å)		Mean Deviations	
Set	Model	Algorithm	All	Side-Chain	Bond (Å)	Angle (^{o})
3F1W	PDB_Redo		0.80	1.12	0.011	2.5
	OPLS-AA/L	Minimize	0.55	0.72	0.010	2.4
		+ DEE	0.81	1.12	0.011	2.4
	AMOEBA	Minimize	0.47	0.61	0.014	2.7
		+ DEE	0.65	0.89	0.013	2.7
3GPM	PDB_Redo		0.47	0.58	0.012	2.5
	OPLS-AA/L	Minimize	1.10	1.40	0.010	2.4
		+ DEE	1.22	1.58	0.010	2.4
	AMOEBA	Minimize	0.98	1.23	0.014	2.7
		+ DEE	1.07	1.36	0.014	2.7
3GPN	PDB_Redo		0.23	0.28	0.012	2.5
	OPLS-AA/L	Minimize	0.63	0.87	0.011	2.3
		+ DEE	0.85	1.19	0.011	2.3
	AMOEBA	Minimize	0.56	0.77	0.014	2.6
		+ DEE	0.74	1.03	0.014	2.6
3L0W	PDB_Redo		0.42	0.46	0.011	2.4
	OPLS-AA/L	Minimize	1.03	1.28	0.009	2.3
		+ DEE	1.20	1.51	0.010	2.3
	AMOEBA	Minimize	0.87	1.05	0.013	2.7
		+ DEE	1.00	1.24	0.012	2.6
3L0X	PDB_Redo		0.29	0.36	0.011	2.5
	OPLS-AA/L	Minimize	0.58	0.73	0.011	2.4
		+ DEE	0.82	1.09	0.011	2.3
	AMOEBA	Minimize	0.48	0.60	0.014	2.7
		+ DEE	0.79	1.07	0.013	2.7
3L10	PDB_Redo		0.45	0.52	0.011	2.4
	OPLS-AA/L	Minimize	1.05	1.30	0.009	2.3
		+ DEE	1.24	1.58	0.010	2.4
	AMOEBA	Minimize	0.85	1.03	0.013	2.7
		+ DEE	1.06	1.31	0.013	2.7
WT	PDB_Redo		1.04	1.42	0.015	3.3
	OPLS-AA/L	Minimize	0.60	0.63	0.011	2.5
		+ DEE	0.95	1.26	0.011	2.5
	AMOEBA	Minimize	0.52	0.63	0.014	2.7
		+ DEE	0.91	1.22	0.014	2.7
Mean	PDB_Redo		0.53	0.68	0.012	2.6
	OPLS-AA/L	Minimize	0.79	0.99	0.010	2.4
		+ DEE	1.03	1.36	0.010	2.4
	AMOEBA	Minimize	0.67	0.84	0.014	2.7
		+ DEE	0.89	1.16	0.013	2.7

Table S3. A comparison of structure quality metrics after AMOEBA DEE refinement using 2-body and 3-body approximations is shown. The AMOEBA 2-body optimization provides higher quality structures than OPLS-AA/L (see Table 2), however, 3-body optimization yields additional improvements beyond those of all other strategies.

PDB					MolProbity		Clash		Ramachandran		Poor
Res.	Refinement	R	R_{free}	E_{FF}	Score	%	Score	%	Out. %	Fav. %	Rotamers %
3f1w	Original	23.46	25.87		2.81	81	35.3	65	0.4	95.2	3.9
2.9 Å	2-body	22.74	27.02	-257	1.56	100	1.0	100	0.8	94.1	3.0
	3-body	22.12	26.25	-129	1.03	100	1.2	100	0.4	96.4	0.4
3gpm	Original	35.42	34.31		3.43	73	52.9	51	4.0	89.3	7.5
3.8 Å	2-body	24.13	27.20	-172	1.41	100	0.0	100	4.8	85.3	2.2
	3-body	24.42	27.25	-110	1.33	100	0.0	100	4.4	85.7	1.8
3gpn	Original	23.81	27.29		2.19	91	11.8	92	0.0	98.0	6.2
2.5 Å	2-body	20.97	25.55	-307	1.58	100	1.8	100	0.0	96.0	3.1
	3-body	20.98	25.59	-351	1.28	100	0.5	100	0.0	96.4	3.1
3l0w	Original	31.45	33.17		3.57	23	51.0	20	0.0	92.3	15.8
2.8 Å	2-body	27.10	29.66	-140	1.34	100	0.4	100	0.6	94.4	2.7
	3-body	27.12	29.60	-260	1.17	100	1.0	100	0.6	94.7	1.7
3l0x	Original	24.27	25.65		2.79	86	15.0	97	0.0	94.4	9.2
3.0 Å	2-body	20.70	24.47	-117	1.37	100	1.7	100	2.0	93.2	0.4
	3-body	20.98	24.43	-125	1.25	100	1.8	100	1.6	92.8	0.4
3l10	Original	31.83	34.36		3.56	23	51.8	20	0.0	92.3	15.1
2.8 Å	2-body	26.88	29.89	-319	1.69	100	0.8	100	0.6	92.6	4.4
	3-body	26.82	29.56	-360	1.59	100	1.0	100	0.6	91.6	3.4
WT	Original	24.89	27.25		1.65	100	5.5	100	0.4	94.9	0.9
3.0 Å	2-body	22.29	24.34	-665	1.15	100	1.0	100	1.2	94.9	0.9
	3-body	21.63	24.24	-760	1.09	100	0.7	100	1.2	94.9	0.9
Mean	Original	27.88	29.70		2.86	68	31.9	64	0.7	93.8	8.4
3.0 Å	2-body	23.54	26.88	-282	1.44	100	0.94	100	1.4	92.9	2.4
	3-body	23.44	26.70	-299	1.25	100	0.88	100	1.3	93.2	1.7
	Δ 3- vs. 2-body	-0.10	-0.18	-17	-0.19	0	-0.06	0	-0.2	0.3	-0.7

Table S4. Run times by structure and method on a single 16-core compute node at 2.6GHz. OPLS-AA/L and AMOEBA are nearly equivalent for minimization due to the X-ray scattering term being the limiting factor. The cost of AMOEBA 3-body DEE is approximately 15x greater than OPLS-AA/L 2-body DEE due to 1) the increased cost of each energy evaluation and 2) computation of 3-body terms.

PDB ID	Force Field	Algorithm	Run Time
3f1w	OPLS-AA/L	Minimize	35 sec
		+ DEE	25 hours
	AMOEBA	Minimize	39 sec
		+ DEE	17 days
3gpm	OPLS-AA/L	Minimize	29 sec
		+ DEE	18 hours
	AMOEBA	Minimize	30 sec
		+ DEE	18 days
3gpn	OPLS-AA/L	Minimize	23 sec
		+ DEE	11 hours
	AMOEBA	Minimize	28 sec
		+ DEE	12 days
3l0w	OPLS-AA/L	Minimize	43 sec
		+ DEE	28 hours
	AMOEBA	Minimize	51 sec
		+ DEE	15 days
3l0x	OPLS-AA/L	Minimize	30 sec
		+ DEE	18 hours
	AMOEBA	Minimize	39 sec
		+ DEE	11 days
3l10	OPLS-AA/L	Minimize	80 sec
		+ DEE	29 hours
	AMOEBA	Minimize	53 sec
		+ DEE	16 days
WT	OPLS-AA/L	Minimize	30 sec
		+ DEE	20 hours
	AMOEBA	Minimize	36 sec
		+ DEE	20 days
Mean	OPLS-AA/L	Minimize	39 sec
		+ DEE	21 hours
	AMOEBA	Minimize	39 sec
		+ DEE	15 days

Supplemental Derivations

I. Many-Body Inclusive Singles Elimination Criterion

We start from the knowledge that any given point in the global rotamer space has a minimum energy equal to the global minimum energy conformation.

$$E_{\text{global}} \geq E_{\text{GMEC}}$$

Equation 1

We denote a superscript g as being the rotamer of a particular residue as it exists in the global minimum energy conformation.

$$E_{\text{global}} = E_{BB} + E(r_i^{\alpha}) + \sum_{j\prime}^{n}\{E(r_j^{g}) + \sum_{k\prime}^{n}[E(r_j^{g}, r_k^{g}) + \sum_{l\prime}^{n}(E(r_j^{g}, r_k^{g}, r_l^{g}) + \cdots)]\} + \sum_{j\prime}^{n}\{E(r_i^{\alpha}, r_j^{g}) + \sum_{k\prime}^{n}[E(r_i^{\alpha}, r_j^{g}, r_k^{g}) + \sum_{l\prime}^{n}(E(r_i^{\alpha}, r_j^{g}, r_k^{g}, r_l^{g}) + \cdots)]\}$$

Equation 2

$$E_{\text{GMEC}} = E_{BB} + E(r_i^{g}) + \sum_{j\prime}^{n}\{E(r_j^{g}) + \sum_{k\prime}^{n}[E(r_j^{g}, r_k^{g}) + \sum_{l\prime}^{n}(E(r_j^{g}, r_k^{g}, r_l^{g}) + \cdots)]\} + \sum_{j\prime}^{n}\{E(r_i^{g}, r_j^{g}) + \sum_{k\prime}^{n}[E(r_i^{g}, r_j^{g}, r_k^{g}) + \sum_{l\prime}^{n}(E(r_i^{g}, r_j^{g}, r_k^{g}, r_l^{g}) + \cdots)]\}$$

Equation 3

Herein E_{BB} is the backbone energy, $E(r_i^{\alpha})$ is the self-energy of residue i in rotamer α, $E\left(r_i^{\alpha}, r_j^{\beta}\right)$ is the two-body energy of residues i, j in rotamers α, β and so on. Self, two-body, and many-body energies are as defined in the main text. Ellipses signify the presence of higher-order terms out to n-body, where n is the number of residues in the system. After explicitly enumerating all energy components including many-body energy, we substitute Eqs. 2 and 3 into Equation 1. Terms without dependence on r_i cancel out. We then find an expression for the remaining portion that doesn't require knowledge of the GMEC conformation.

$$\begin{cases} \max\limits_{s}\left[E(r_i^{\alpha}, r_k^{s}) + \sum\limits_{l\prime}^{n} E(r_i^{\alpha}, r_k^{s}, r_l^{g}) + \cdots\right] \geq E(r_i^{\alpha}, r_k^{g}) + \sum\limits_{l'}^{n} E(r_i^{\alpha}, r_k^{g}, r_l^{g}) + \cdots \\ \min\limits_{s}\left[E(r_i^{g}, r_k^{s}) + \sum\limits_{l\prime}^{n} E(r_i^{g}, r_k^{s}, r_l^{g}) + \cdots\right] \leq E(r_i^{g}, r_k^{g}) + \sum\limits_{l'}^{n} E(r_i^{g}, r_k^{g}, r_l^{g}) + \cdots \end{cases}$$

Equation 4

$$\begin{cases} \max\limits_{t}[E(r_i^{\alpha}, r_k^{s}, r_l^{t}) + \cdots] \geq E\left(r_i^{\alpha}, r_k^{s}, r_l^{g}\right) + \cdots \\ \min\limits_{t}\left[E\left(r_i^{g}, r_k^{s}, r_l^{t}\right) + \cdots\right] \leq E\left(r_i^{g}, r_k^{s}, r_l^{g}\right) + \cdots \end{cases}$$

Equation 5

$$\begin{cases} \max\limits_{s}\left[E(r_i^{\alpha}, r_k^{s}) + \sum\limits_{l\prime}^{n} \max\limits_{t}\left(E(r_i^{\alpha}, r_k^{s}, r_l^{t})\right) + \cdots\right] \geq E\left(r_i^{\alpha}, r_k^{g}\right) + \sum\limits_{l'}^{n} E\left(r_i^{\alpha}, r_k^{g}, r_l^{g}\right) + \cdots \\ \min\limits_{s}\left[E\left(r_i^{g}, r_k^{s}\right) + \sum\limits_{l\prime}^{n} \min\limits_{t}\left(E\left(r_i^{g}, r_k^{s}, r_l^{t}\right)\right) + \cdots\right] \leq E\left(r_i^{g}, r_k^{g}\right) + \sum\limits_{l'}^{n} E\left(r_i^{g}, r_k^{g}, r_l^{g}\right) + \cdots \end{cases}$$

Equation 6

Expressing the substituted Eq. 1 using the left-hand side of Eq. 6 yields the final singles elimination criterion.

II. Many-Body Inclusive Pairwise Elimination Criterion

$$E_{\text{global}} \geq E_{\text{GMEC}}$$

Equation 7

$$\begin{aligned} &E(r_i^{\alpha}) + E\left(r_j^{\beta}\right) + E\left(r_i^{\alpha}, r_j^{\beta}\right) + \textstyle\sum_{k\prime}^{n}\left\{E\left(r_i^{\alpha}, r_k^{g}\right) + E\left(r_j^{\beta}, r_k^{g}\right) + E\left(r_i^{\alpha}, r_j^{\beta}, r_k^{g}\right) + \right. \\ &\left.\textstyle\sum_{l\prime}^{n}\left[E\left(r_i^{\alpha}, r_k^{g}, r_l^{g}\right) + E\left(r_j^{\beta}, r_k^{g}, r_l^{g}\right) + E\left(r_i^{\alpha}, r_j^{\beta}, r_k^{g}, r_l^{g}\right) \ldots\right]\right\} \geq \\ &E\left(r_i^{g}\right) + E\left(r_j^{g}\right) + E\left(r_i^{g}, r_j^{g}\right) + \textstyle\sum_{k\prime}^{n}\left\{E\left(r_i^{g}, r_k^{g}\right) + E\left(r_j^{g}, r_k^{g}\right) + E\left(r_i^{g}, r_j^{g}, r_k^{g}\right) + \right. \\ &\left.\textstyle\sum_{l\prime}^{n}\left[E\left(r_i^{g}, r_k^{g}, r_l^{g}\right) + E\left(r_j^{g}, r_k^{g}, r_l^{g}\right) + E\left(r_i^{g}, r_j^{g}, r_k^{g}, r_l^{g}\right) \ldots\right]\right\} \end{aligned}$$

Equation 8

We begin again from Equation 7. After explicitly enumerating all energy components of E_{global}, E_{GMEC} and substituting, all terms not involving r_i or r_j cancel out. We then find expressions for the remaining terms that do not require knowledge of the GMEC configuration.

$$\max_s \Bigg[E(r_i^\alpha, r_k^s) + E\left(r_j^\beta, r_k^s\right) + E\left(r_i^\alpha, r_j^\beta, r_k^s\right) + \sum_{l'}^{n} \left(E\left(r_i^\alpha, r_k^s, r_l^g\right) + E\left(r_j^\beta, r_k^s, r_l^g\right) + E\left(r_i^\alpha, r_j^\beta, r_k^s, r_l^g\right) + \cdots \right) \Bigg] \geq E\left(r_i^\alpha, r_k^g\right) + E\left(r_j^\beta, r_k^g\right) + E\left(r_i^\alpha, r_j^\beta, r_k^g\right) + \sum_{l'}^{n} \left(E\left(r_i^\alpha, r_k^g, r_l^g\right) + E\left(r_j^\beta, r_k^g, r_l^g\right) + E\left(r_i^\alpha, r_j^\beta, r_k^g, r_l^g\right) + \cdots \right)$$

Equation 9

$$\min_s \Bigg[E\left(r_i^g, r_k^s\right) + E\left(r_j^g, r_k^s\right) + E\left(r_i^g, r_j^g, r_k^s\right) + \sum_{l'}^{n} \left(E\left(r_i^g, r_k^s, r_l^g\right) + E\left(r_j^g, r_k^s, r_l^g\right) + E\left(r_i^g, r_j^g, r_k^s, r_l^g\right) + \cdots \right) \Bigg] \leq E\left(r_i^g, r_k^g\right) + E\left(r_j^g, r_k^g\right) + E\left(r_i^g, r_j^g, r_k^g\right) + \sum_{l'}^{n} \left(E\left(r_i^g, r_k^g, r_l^g\right) + E\left(r_j^g, r_k^g, r_l^g\right) + E\left(r_i^g, r_j^g, r_k^g, r_l^g\right) + \cdots \right)$$

Equation 10

$$\max_t \left[E(r_i^\alpha, r_k^s, r_l^t) + E\left(r_j^\beta, r_k^s, r_l^t\right) + E\left(r_i^\alpha, r_j^\beta, r_k^s, r_l^t\right) + \cdots \right] \geq E\left(r_i^\alpha, r_k^s, r_l^g\right) + E\left(r_j^\beta, r_k^s, r_l^g\right) + E\left(r_i^\alpha, r_j^\beta, r_k^s, r_l^g\right) + \cdots$$

$$\min_t \left[E\left(r_i^g, r_k^s, r_l^t\right) + E\left(r_j^g, r_k^s, r_l^t\right) + E\left(r_i^g, r_j^g, r_k^s, r_l^t\right) + \cdots \right] \leq E\left(r_i^g, r_k^s, r_l^g\right) + E\left(r_j^g, r_k^s, r_l^g\right) + E\left(r_i^g, r_j^g, r_k^s, r_l^g\right) + \cdots$$

Equation 11

Substituting Eq. 11 into Eqs. 9 and 10, we get:

$$\max_s \Bigg[E(r_i^\alpha, r_k^s) + E\left(r_j^\beta, r_k^s\right) + E\left(r_i^\alpha, r_j^\beta, r_k^s\right) + \sum_{l'}^{n} \max_t \left(E(r_i^\alpha, r_k^s, r_l^t) + E\left(r_j^\beta, r_k^s, r_l^t\right) + E\left(r_i^\alpha, r_j^\beta, r_k^s, r_l^t\right) + \cdots \right) \Bigg] \geq E\left(r_i^\alpha, r_k^g\right) + E\left(r_j^\beta, r_k^g\right) + E\left(r_i^\alpha, r_j^\beta, r_k^g\right) + \sum_{l'}^{n} E\left(r_i^\alpha, r_k^g, r_l^g\right) + E\left(r_j^\beta, r_k^g, r_l^g\right) + E\left(r_i^\alpha, r_j^\beta, r_k^g, r_l^g\right) + \cdots$$

Equation 12

$$\min_s \Bigg[E\left(r_i^g, r_k^s\right) + E\left(r_j^g, r_k^s\right) + E\left(r_i^g, r_j^g, r_k^s\right) + \sum_{l'}^{n} \min_t \left(E\left(r_i^g, r_k^s, r_l^t\right) + E\left(r_j^g, r_k^s, r_l^t\right) + E\left(r_i^g, r_j^g, r_k^s, r_l^t\right) + \cdots \right) \Bigg]$$
$$\leq E\left(r_i^g, r_k^g\right) + E\left(r_j^g, r_k^g\right) + E\left(r_i^g, r_j^g, r_k^g\right) + \sum_{l'}^{n} E\left(r_i^g, r_k^g, r_l^g\right) + E\left(r_j^g, r_k^g, r_l^g\right) + E\left(r_i^g, r_j^g, r_k^g, r_l^g\right) + \cdots$$

Equation 13

Expressing Equation 7 using the left-hand side of Eqs. 12 and 13 (for E_{global} and E_{GMEC} respectively) yields the final pairwise elimination criterion.

III. Many-Body Generalized Goldstein Singles Elimination Criterion

We begin from the substituted Eqs 1 through 3.

$$E(r_i^\alpha) + \sum_{j'}^{n} \left\{ E\left(r_i^\alpha, r_j^g\right) + \sum_{k'}^{n} \left[E\left(r_i^\alpha, r_j^g, r_k^g\right) + \sum_{l'}^{n} \left(E\left(r_i^\alpha, r_j^g, r_k^g, r_l^g\right) + \cdots \right) \right] \right\} \geq$$
$$E\left(r_i^g\right) + \sum_{j'}^{n} \left\{ E\left(r_i^g, r_j^g\right) + \sum_{k'}^{n} \left[E\left(r_i^g, r_j^g, r_k^g\right) + \sum_{l'}^{n} \left(E\left(r_i^g, r_j^g, r_k^g, r_l^g\right) + \cdots \right) \right] \right\}$$

Equation 14

In contrast to the original singles elimination derivation, we first subtract the right-hand side before applying the min operator.

$$E(r_i^\alpha) - E\left(r_i^g\right) + \sum_{j'}^{n} \left\{ E\left(r_i^\alpha, r_j^g\right) - E\left(r_i^g, r_j^g\right) + \sum_{k'}^{n} \left[E\left(r_i^\alpha, r_j^g, r_k^g\right) - E\left(r_i^g, r_j^g, r_k^g\right) + \right.\right.$$
$$\left.\left. \sum_{l'}^{n} \left(E\left(r_i^\alpha, r_j^g, r_k^g, r_l^g\right) - E\left(r_i^g, r_j^g, r_k^g, r_l^g\right) \ldots \right) \right] \right\} \geq 0$$

Equation 15

$$\min_s \Bigg\{ E\left(r_i^{\alpha}, r_j^{s}\right) - E\left(r_i^{g}, r_j^{s}\right) + \sum_{k'}^{n} \Bigg[E\left(r_i^{\alpha}, r_j^{s}, r_k^{g}\right) - E\left(r_i^{g}, r_j^{s}, r_k^{g}\right) + \sum_{l'}^{n} \left(E\left(r_i^{\alpha}, r_j^{s}, r_k^{g}, r_l^{g}\right) - E\left(r_i^{g}, r_j^{s}, r_k^{g}, r_l^{g}\right) \dots \right) \Bigg] \Bigg\} \leq E\left(r_i^{\alpha}, r_j^{g}\right) - E\left(r_i^{g}, r_j^{g}\right) + \sum_{k'}^{n} \Bigg[E\left(r_i^{\alpha}, r_j^{g}, r_k^{g}\right) - E\left(r_i^{g}, r_j^{g}, r_k^{g}\right) + \sum_{l'}^{n} \left(E\left(r_i^{\alpha}, r_j^{g}, r_k^{g}, r_l^{g}\right) - E\left(r_i^{g}, r_j^{g}, r_k^{g}, r_l^{g}\right) \dots \right) \Bigg]$$

Equation 16

$$\min_t \left[E\left(r_i^{\alpha}, r_j^{s}, r_k^{t}\right) - E\left(r_i^{g}, r_j^{s}, r_k^{t}\right) + \sum_{l'}^{n} \left(E\left(r_i^{\alpha}, r_j^{s}, r_k^{t}, r_l^{g}\right) - E\left(r_i^{g}, r_j^{s}, r_k^{t}, r_l^{g}\right) \dots \right) \right] \leq E\left(r_i^{\alpha}, r_j^{s}, r_k^{g}\right) - E\left(r_i^{g}, r_j^{s}, r_k^{g}\right) + \sum_{l'}^{n} \left(E\left(r_i^{\alpha}, r_j^{s}, r_k^{g}, r_l^{g}\right) - E\left(r_i^{g}, r_j^{s}, r_k^{g}, r_l^{g}\right) \dots \right)$$

Equation 17

$$\min_u \left(E\left(r_i^{\alpha}, r_j^{s}, r_k^{t}, r_l^{u}\right) - E\left(r_i^{g}, r_j^{s}, r_k^{t}, r_l^{u}\right) \dots \right) \leq \left(E\left(r_i^{\alpha}, r_j^{s}, r_k^{t}, r_l^{g}\right) - E\left(r_i^{g}, r_j^{s}, r_k^{t}, r_l^{g}\right) \dots \right)$$

Equation 18

As before, we then identify max and min inequalities that relieve us of reliance on knowing g. Substituting Eqs. 16-18 into Eq. 15, we arrive at the general Goldstein criterion.

IV. Many-Body Generalized Goldstein Pairwise Elimination Criterion

This derivation follows from the many-body Goldstein singles elimination in the same fashion that the original pairwise elimination followed from the original singles elimination.

$$E(r_i^{\alpha}) + E\left(r_j^{\beta}\right) + E\left(r_i^{\alpha}, r_j^{\beta}\right) + \sum_{k'}^{n} \left\{ E\left(r_i^{\alpha}, r_k^{g}\right) + E\left(r_j^{\beta}, r_k^{g}\right) + E\left(r_i^{\alpha}, r_j^{\beta}, r_k^{g}\right) + \sum_{l'}^{n} \left[E\left(r_i^{\alpha}, r_k^{g}, r_l^{g}\right) + E\left(r_j^{\beta}, r_k^{g}, r_l^{g}\right) + E\left(r_i^{\alpha}, r_j^{\beta}, r_k^{g}, r_l^{g}\right) + \cdots \right] \right\} \geq$$

$$E\left(r_i^g\right) + E\left(r_j^g\right) + E\left(r_i^g, r_j^g\right) + \sum_{k\prime}^{n}\{E\left(r_i^g, r_k^g\right) + E\left(r_j^g, r_k^g\right) + E\left(r_i^g, r_j^g, r_k^g\right) + \sum_{l\prime}^{n}\left[E\left(r_i^g, r_k^g, r_l^g\right) + E\left(r_j^g, r_k^g, r_l^g\right) + E\left(r_i^g, r_j^g, r_k^g, r_l^g\right) + \cdots\right]\}$$

Equation 19

$$E(r_i^\alpha) - E\left(r_i^g\right) + E\left(r_j^\beta\right) - E\left(r_j^g\right) + E\left(r_i^\alpha, r_j^\beta\right) - E\left(r_i^g, r_j^g\right) + \sum_{k\prime}^{n}\left\{E\left(r_i^\alpha, r_k^g\right) - E\left(r_i^g, r_k^g\right) + E\left(r_j^\beta, r_k^g\right) - E\left(r_j^g, r_k^g\right) + E\left(r_i^\alpha, r_j^\beta, r_k^g\right) - E\left(r_i^g, r_j^g, r_k^g\right) + \sum_{l\prime}^{n}\left[E\left(r_i^\alpha, r_k^g, r_l^g\right) - E\left(r_i^g, r_k^g, r_l^g\right) + E\left(r_j^\beta, r_k^g, r_l^g\right) - E\left(r_j^g, r_k^g, r_l^g\right) + E\left(r_i^\alpha, r_j^\beta, r_k^g, r_l^g\right) - E\left(r_i^g, r_j^g, r_k^g, r_l^g\right) + \cdots\right]\right\} \geq 0$$

Equation 20

$$\min_s\left[E(r_i^\alpha, r_k^s) - E\left(r_i^g, r_k^s\right) + E\left(r_j^\beta, r_k^s\right) - E\left(r_j^g, r_k^s\right) + E\left(r_i^\alpha, r_j^\beta, r_k^s\right) + E\left(r_i^g, r_j^g, r_k^s\right) + \cdots\right] \leq E\left(r_i^\alpha, r_k^g\right) - E\left(r_i^g, r_k^g\right) + E\left(r_j^\beta, r_k^g\right) - E\left(r_j^g, r_k^g\right) + E\left(r_i^\alpha, r_j^\beta, r_k^g\right) + E\left(r_i^g, r_j^g, r_k^g\right) + \cdots$$

Equation 21

Downstream min and max operators are applied just as before and are substituted into Eq. 20 to yield the many-body generalized pairwise Goldstein criterion.