nature portfolio

https://doi.org/10.1038/s42256-021-00366-3

Supplementary information

Dynamic particle swarm optimization of biomolecular simulation parameters with flexible objective functions

In the format provided by the authors and unedited

nature portfolio

https://doi.org/10.1038/s42256-021-00366-3

Supplementary information

Dynamic particle swarm optimization of biomolecular simulation parameters with flexible objective functions

In the format provided by the authors and unedited

Exhaustive FLAPS optimization results

Seed	1790954	1791103	1791104	1791105	1791106	1792508	1792509
ρ^{RMSD}	0.85	0.70	0.88	0.63	0.84	0.64	0.79
ρ^{GDT}	-0.94	-0.84	-0.87	-0.68	-0.87	-0.74	-0.80
f^{min}	-2.34	-2.03	-1.79	-2.15	-1.99	-1.60	-1.78
f^{max}	8.32	8.31	5.86	4.38	4.41	4.72	6.27
Best simulation in terms of OF							
k_χ	2.170 e-10	1.073 e-10	3.339 e-11	6.654 e-11	5.080 e-11	1.493 e-09	5.869 e-11
T	13.19	10.36	28.82	10.47	11.05	13.60	14.65
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.1	2.1	2.2	2.2	2.2	2.9	2.2
$\mathrm{GDT}^{\mathrm{med}}$	70.59	69.96	69.22	69.44	69.44	64.81	69.22
Best simulation in terms of $\mathrm{RMSD}^{\mathrm{med}}$							
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.1	2.1	2.2	2.1	2.1	2.2	2.1
$f\left(\mathrm{RMSD}^{\mathrm{med}}\right)$	-2.03	-2.02	-1.73	-1.79	-1.55	-0.90	-1.45
k_χ	3.001 e-10	2.320 e-10	3.422 e-11	2.401 e-10	4.190 e-10	5.652 e-10	2.395 e-10
T	11.98	10.43	29.63	10.18	10.03	23.67	12.93
$\mathrm{GDT}^{\mathrm{med}}$	70.69	70.80	69.54	70.91	70.69	68.17	70.69
Best simulation in terms of $\mathrm{GDT}^{\mathrm{med}}$							
$\mathrm{GDT}^{\mathrm{med}}$	70.69	70.80	69.54	70.91	70.69	68.17	70.69
$f\left(\mathrm{GDT}^{\mathrm{med}}\right)$	-2.03	-2.02	-1.73	-1.79	-1.55	-1.08	-1.45
k_χ	3.001 e-10	2.320 e-10	3.422 e-11	2.401 e-10	4.190 e-10	6.429 e-10	2.395 e-10
T	11.98	10.43	29.63	10.18	10.03	20.89	12.93
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.1	2.1	2.2	2.1	2.1	2.2	2.1

Table S1. FLAPS optimization results for SAXS-guided holo-to-apo transition of LAO protein. OF, objective function f, ρ^{RMSD}, Pearson correlation of OF and $\mathrm{RMSD}^{\mathrm{med}}$, ρ^{GDT}, Pearson correlation of OF and $\mathrm{GDT}^{\mathrm{med}}$, k_χ, bias weight, T, temperature, $\mathrm{RMSD}^{\mathrm{med}}$, median root-mean-square deviation, $\mathrm{GDT}^{\mathrm{med}}$, median global distance test (total score).

Seed	1795691	1797335	1797338	1797339	1798723	1801054	1810891
ρ^{RMSD}	0.58	0.33	0.40	0.41	0.49	0.03	0.42
ρ^{GDT}	-0.85	-0.61	-0.77	-0.76	-0.81	-0.32	-0.74
f^{min}	-1.42	-1.54	-1.20	-1.50	-1.57	-1.44	-1.62
f^{max}	6.92	6.38	7.09	6.63	9.05	5.45	7.47
Best simulation in terms of OF							
k_χ	1.969 e-09	1.283 e-09	1.858 e-09	1.810 e-09	1.970 e-09	8.218 e-10	1.819 e-09
T	16.90	10.02	10.17	10.28	10.56	10.51	10.09
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.8	3.1	2.9	2.9	2.8	3.3	2.9
$\mathrm{GDT}^{\mathrm{med}}$	63.20	62.38	63.67	63.55	63.78	60.63	63.55
Best simulation in terms of $\mathrm{RMSD}^{\mathrm{med}}$							
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.6	2.6	2.6	2.6	2.6	2.7	2.6
$f\left(\mathrm{RMSD}^{\mathrm{med}}\right)$	-1.19	-0.46	-0.59	-0.81	-1.21	2.15	-1.28
k_χ	2.910 e-09	2.996 e-09	3.477 e-09	2.488 e-09	3.213 e-09	3.311 e-09	3.419 e-09
T	12.40	16.32	12.02	15.51	10.52	31.12	10.01
$\mathrm{GDT}^{\mathrm{med}}$	63.20	62.97	63.20	63.09	63.44	62.38	63.32
Best simulation in terms of $\mathrm{GDT}^{\mathrm{med}}$							
$\mathrm{GDT}^{\mathrm{med}}$	63.78	63.90	63.90	63.90	63.78	62.62	63.90
$f\left(\mathrm{GDT}^{\mathrm{med}}\right)$	-1.28	-1.29	-1.17	-1.35	-1.57	-1.34	-1.59
k_χ	2.030 e-09	2.171 e-09	2.134 e-09	2.171 e-09	1.970 e-09	1.377 e-09	2.170 e-09
T	10.84	10.67	10.07	10.45	10.56	10.22	10.23
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.8	2.8	2.8	2.8	2.8	3.0	2.8

Table S2. FLAPS optimization results for SAXS-guided open-to-closed transition of adenylate kinase. OF, objective function f, ρ^{RMSD}, Pearson correlation of OF and $\mathrm{RMSD}^{\mathrm{med}}$, ρ^{GDT}, Pearson correlation of OF and $\mathrm{GDT}^{\mathrm{med}}$, k_χ, bias weight, T, temperature, $\mathrm{RMSD}^{\mathrm{med}}$, median root-mean-square deviation, $\mathrm{GDT}^{\mathrm{med}}$, median global distance test (total score).

Seed	1800990	1800994	1800995	1800996	1805228	1805229	1805230
ρ^{RMSD}	0.78	0.88	0.83	0.84	0.69	0.87	0.59
ρ^{GDT}	-0.75	-0.86	-0.80	-0.81	-0.67	-0.88	-0.56
f^{min}	-1.33	-1.34	-1.27	-1.51	-1.82	-1.65	-1.58
f^{max}	6.28	5.45	6.76	6.00	3.59	5.53	5.25
Best simulation in terms of OF							
k_χ	1.035 e-09	1.245 e-09	9.913 e-10	9.194 e-10	6.466 e-10	1.252 e-09	7.810 e-10
T	10.25	10.53	10.46	10.01	10.74	10.90	13.23
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.2	2.0	2.2	2.2	2.3	2.0	2.2
$\mathrm{GDT}^{\mathrm{med}}$	67.02	69.33	66.91	66.81	65.44	69.33	66.17
Best simulation in terms of $\mathrm{RMSD}^{\mathrm{med}}$							
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.0	2.0	2.0	2.0	2.0	2.0	2.0
$f\left(\mathrm{RMSD}^{\mathrm{med}}\right)$	-1.13	-1.24	-1.02	-1.37	-1.50	-1.50	-0.82
k_χ	1.448 e-09	1.473 e-09	1.388 e-09	1.399 e-09	1.503 e-09	1.580 e-09	1.389 e-09
T	11.94	11.66	11.65	10.67	12.74	11.76	17.35
$\mathrm{GDT}^{\mathrm{med}}$	69.44	69.44	69.44	69.44	69.33	69.44	69.22
Best simulation in terms of $\mathrm{GDT}^{\mathrm{med}}$							
$\mathrm{GDT}^{\mathrm{med}}$	69.44	69.44	69.44	69.44	69.33	69.44	69.44
$f\left(\mathrm{GDT}^{\mathrm{med}}\right)$	-1.13	-1.25	-0.97	-1.48	-1.41	-1.50	-1.35
k_χ	1.448 e-09	1.285 e-09	1.328 e-09	1.070 e-09	1.386 e-09	1.580 e-09	1.160 e-09
T	11.94	12.62	13.65	10.92	16.06	11.76	12.35
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.0	2.0	2.0	2.0	2.0	2.0	2.0

Table S3. `FLAPS` optimization results for SAXS-guided apo-to-holo transition of LAO protein. OF, objective function f, ρ^{RMSD}, Pearson correlation of OF and $\mathrm{RMSD}^{\mathrm{med}}$, ρ^{GDT}, Pearson correlation of OF and $\mathrm{GDT}^{\mathrm{med}}$, k_χ, bias weight, T, temperature, $\mathrm{RMSD}^{\mathrm{med}}$, median root-mean-square deviation, $\mathrm{GDT}^{\mathrm{med}}$, median global distance test (total score).

Seed	1799347	1799348	1801030	1801031	1801032	1801033	1801034
ρ^{RMSD}	0.41	0.65	0.79	0.62	0.65	0.71	0.56
ρ^{GDT}	0.29	-0.15	-0.53	-0.17	-0.31	-0.29	-0.11
f^{min}	-1.86	-2.50	-1.98	-2.61	-1.84	-2.37	-1.49
f^{max}	4.75	7.42	6.30	6.08	4.76	6.54	5.68
Best simulation in terms of OF							
k_χ	1.071 e-10	3.888 e-11	6.902 e-11	4.917 e-11	3.563 e-11	6.541 e-11	3.634 e-10
T	26.58	41.99	11.82	37.45	45.43	15.06	10.53
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	4.4	4.4	4.3	4.4	4.4	4.3	4.2
$\mathrm{GDT}^{\mathrm{med}}$	43.11	45.33	45.44	44.27	44.57	45.44	43.46
Best simulation in terms of $\mathrm{RMSD}^{\mathrm{med}}$							
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	3.5	3.5	3.5	3.6	3.6	3.5	3.4
$f\left(\mathrm{RMSD}^{\mathrm{med}}\right)$	-0.36	-0.50	0.26	-0.38	-0.52	0.38	-1.00
k_χ	8.490 e-10	8.294 e-10	8.194 e-10	6.801 e-10	6.623 e-10	1.039 e-09	1.117 e-09
T	31.22	34.12	27.42	43.01	44.02	28.40	17.73
$\mathrm{GDT}^{\mathrm{med}}$	49.71	49.65	49.65	48.83	48.95	50.12	49.94
Best simulation in terms of $\mathrm{GDT}^{\mathrm{med}}$							
$\mathrm{GDT}^{\mathrm{med}}$	50.23	49.65	50.47	50.81	50.59	50.35	51.17
$f\left(\mathrm{GDT}^{\mathrm{med}}\right)$	4.75	-0.50	-0.63	0.43	0.34	-0.11	-0.73
k_χ	3.122 e-11	8.294 e-10	3.134 e-11	2.229 e-11	5.289 e-09	3.515 e-10	2.163 e-09
T	22.77	34.12	13.00	42.58	27.35	30.33	13.43
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	4.9	3.5	4.9	4.9	5.3	4.0	3.5

Table S4. `FLAPS` optimization results for SAXS-guided closed-to-open transition of adenylate kinase. OF, objective function f, ρ^{RMSD}, Pearson correlation of OF and $\mathrm{RMSD}^{\mathrm{med}}$, ρ^{GDT}, Pearson correlation of OF and $\mathrm{GDT}^{\mathrm{med}}$, k_χ, bias weight, T, temperature, $\mathrm{RMSD}^{\mathrm{med}}$, median root-mean-square deviation, $\mathrm{GDT}^{\mathrm{med}}$, median global distance test (total score).

Analyzing swarm convergence

In the context of PSO, convergence can refer to (i) the sequential convergence of solutions, i.e., all particles have reached a particular and possibly (but not necessarily) optimal position, or (ii) convergence of either the personal bests or the global best to a local optimum, independent of the swarm's behavior as a whole. We consider the first convergence concept. Analyzing sequential convergence has issued guidelines for selecting PSO parameters that presumably cause the particles to converge to some point in the search space[19]. A common view is that the swarm varies between exploration and exploitation, which involves adapting the algorithm and its parameters to properly balance these behaviors. This is important to avoid early entrapment in local optima yet allow for a reasonable convergence rate. A primary focus of past research thus was increasing the algorithm's adaptability by making it more complex. However, it is still not clear how the swarm's behavior affects the actual optimization performance, in particular for dynamic environments. This renders the opposite approach, i.e., implementing PSO setups that perform well regardless of how the swarm's behavior is to be interpreted, equally valid. Inspired by Occam's razor, this view is based on the principle that PSO should be simplified to the greatest possible extent without compromising its performance. As a metaheuristic, PSO can only be proven correct in the sense of demonstrating its efficacy empirically by a finite number of computational experiments. This poses the risk of making errors in its description and implementation. Simplifying PSO was first suggested by Kennedy in 1997[17]. The concept has been studied more extensively[26,59], where the optimization performance was found to improve across different problems and the parameters were easier to tune.

Following the "simplifying PSO" paradigm, `FLAPS` rests on a slim PSO core with only few parameters. In the velocity update, cognitive and social influence have random components limited by their associated acceleration coefficients ϕ_1 and ϕ_2 (see Algorithm 1). This randomness keeps the particles from moving directly toward global and personal best. It facilitates exploration of new solutions near the current best positions and diversifies the particles for more effective searches. We find that further diversity enhancement is not needed for the presented application. Supplementary Figs. S1, S2, S3, and S4 show the swarm in the current OF topology after each generation. For all protein systems, the swarm converged to a stable topology and contracted around a functional parameter combination. We calculated a relative swarm spread with respect to the swarm's initial state,

$$SwS = \frac{\mathrm{std}\left(\mathrm{norm}\, T\right) + \mathrm{std}\left(\mathrm{norm}\, k_\chi\right)}{SwS_0}\,, \tag{5}$$

after each generation (Supplementary Figs. S1, S2, S3, and S4, bottom left). $\mathrm{std}\,(x)$ is the standard deviation of quantity x, $\mathrm{norm}\, x$ is the min-max scaled quantity $(x - \min x)\,/\,(\max x - \min x)$, and SwS_0 is the absolute initial swarm spread. Starting from a value of 1 by definition, the swarm spread significantly dropped down in the first two to five generations. Depending on the system, it stabilized at fractions of 0.2 to 0.4 of each initial swarm spread with a tendency to further decrease.

In addition, we considered the Euclidean distance of the current normalized global best from that of each previous generation (Supplementary Figs. S1, S2, S3, and S4, bottom right). This g_{best} fluctuation minimized clearly for only one of the considered systems (Supplementary Fig. S3). As different parameter combinations can equally yield useful results in physico-empirical structure-based models, this is no surprise. Many factors influence the convergence behavior and performance of particle-swarm based algorithms, including selection of the acceleration coefficients, velocity clamping, and the swarm's communication network. As a metaheuristic that implements a form of stochastic optimization, PSO is not guaranteed to find the globally optimal solution. It rather is a practical strategy that guides the optimization process in order to efficiently explore the search space and find near-optimal solutions. While built on the "simplifying PSO" paradigm, `FLAPS` can easily be complemented by concepts such as inertia weight[18] and swarm constriction[19], diversity increasing mechanisms, or flexible termination criteria based on, e.g, the swarm spread or the global best's fluctuation. If desired, the communication pattern can also be adapted towards local geometrical or social topologies.

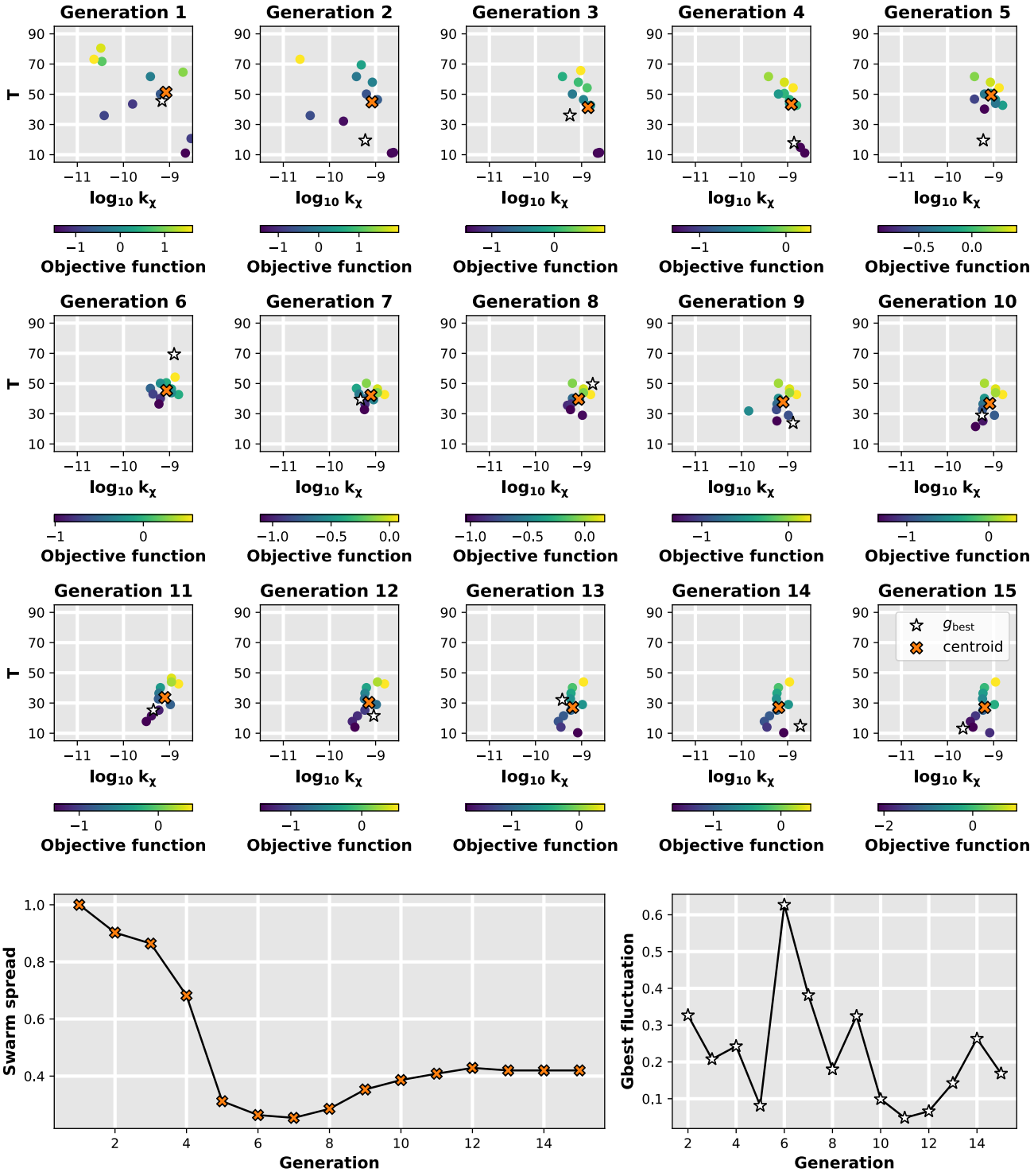

Figure S1. Swarm convergence for SAXS-guided holo-to-apo transition of lysine-, arginine-, ornithine-binding protein (seed 1790954). Top: Dynamically evolving objective-function topology after each generation. The current global best position, g_{best}, and the swarm's centroid are marked. k_χ, bias weight, T, temperature. **Bottom:** Evolution of swarm spread SwS (left) and g_{best} fluctuation (right) during the `FLAPS` optimization.

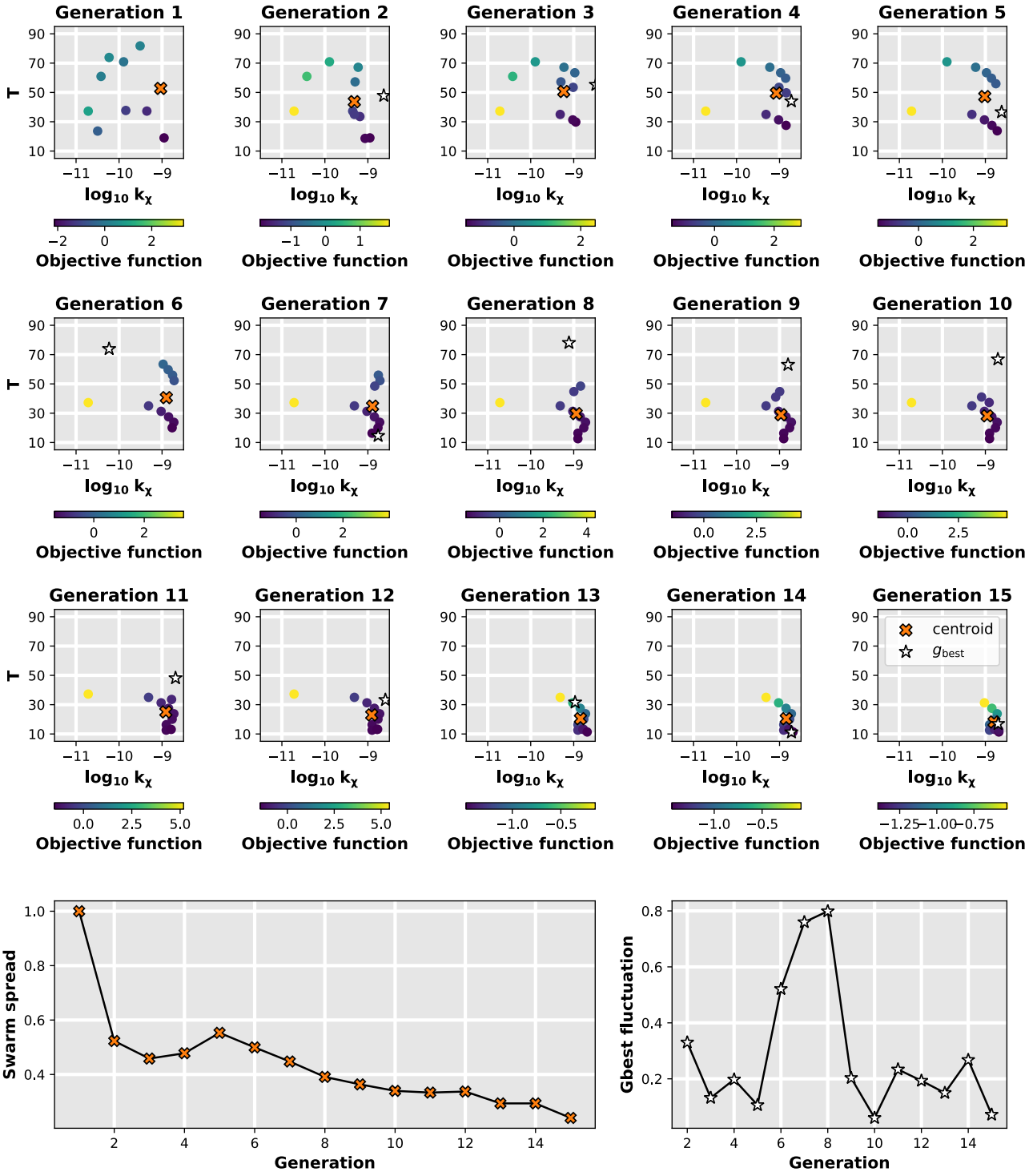

Figure S2. Swarm convergence for SAXS-guided open-to-closed transition of adenylate kinase (seed 1795691). Top: Dynamically evolving objective-function topology after each generation. The current global best position, g_{best}, and the swarm's centroid are marked. k_χ, bias weight, T, temperature. **Bottom:** Evolution of swarm spread SwS (left) and g_{best} fluctuation (right) during the FLAPS optimization.

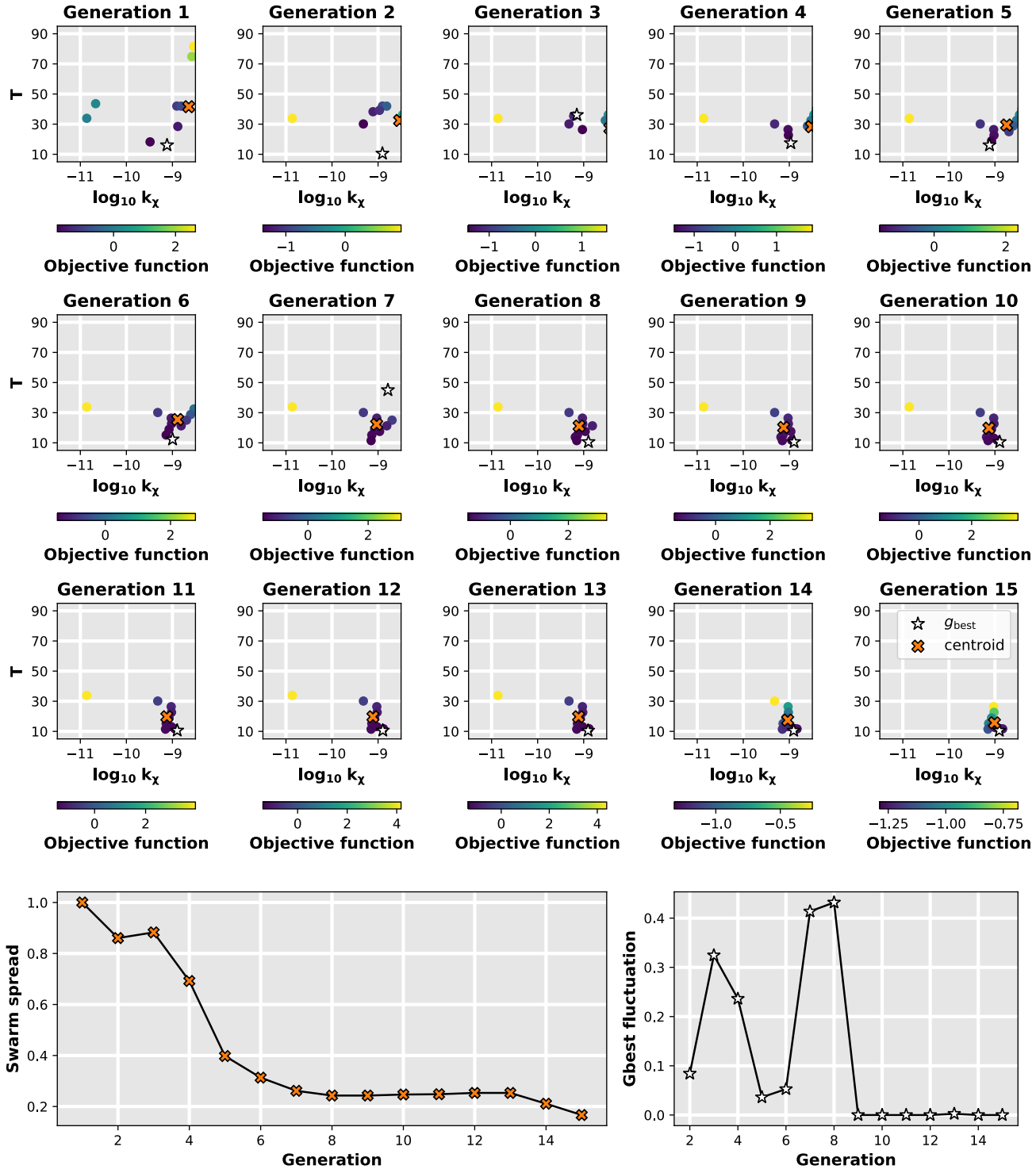

Figure S3. Swarm convergence for SAXS-guided apo-to-holo transition of lysine-, arginine-, ornithine-binding protein (seed 1800994). Top: Dynamically evolving objective-function topology after each generation. The current global best position, g_{best}, and the swarm's centroid are marked. k_χ, bias weight, T, temperature. **Bottom:** Evolution of swarm spread SwS (left) and g_{best} fluctuation (right) during the `FLAPS` optimization.

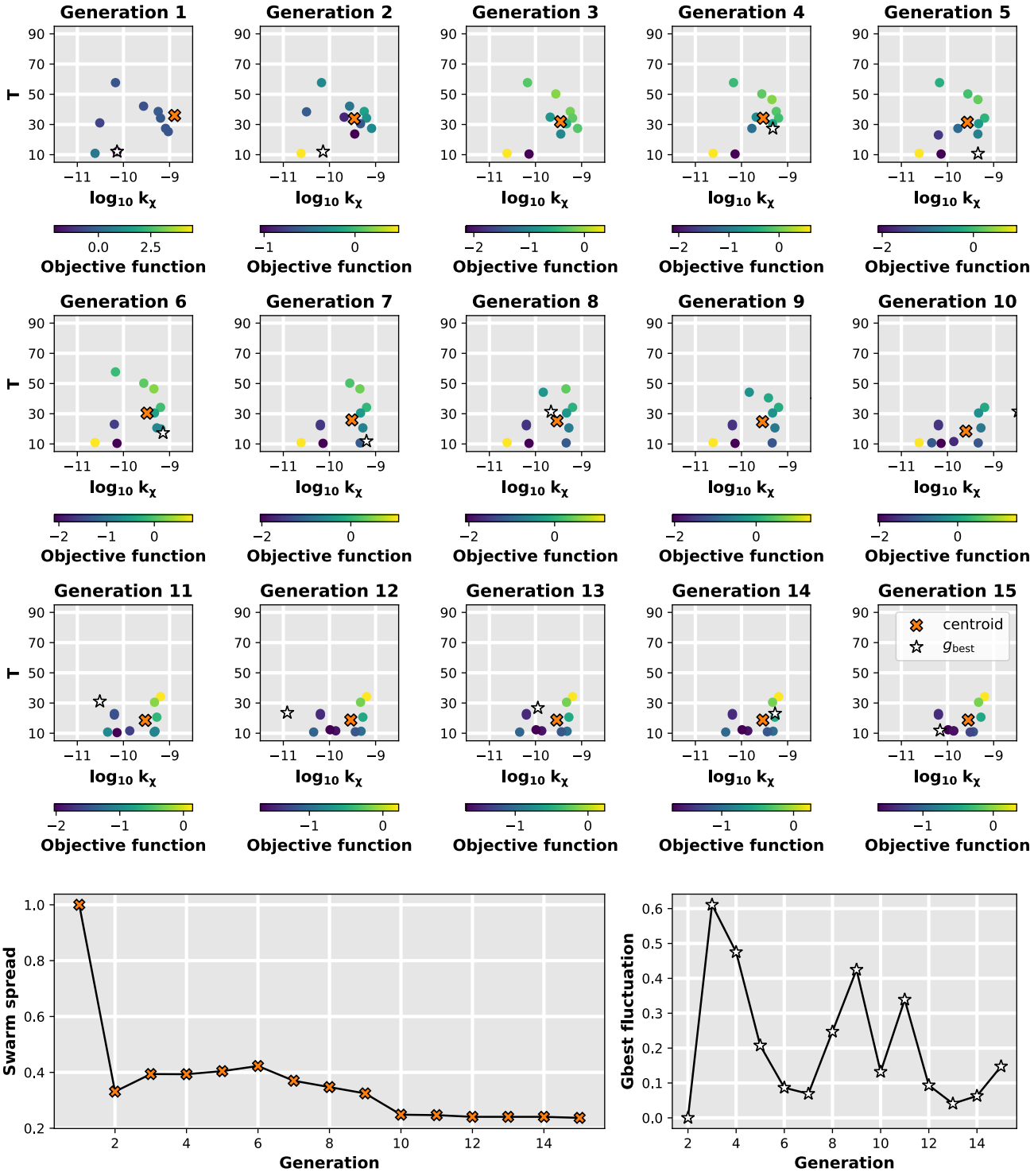

Figure S4. Swarm convergence for SAXS-guided closed-to-open transition of adenylate kinase (seed 1801030). Top: Dynamically evolving objective-function topology after each generation. The current global best position, g_{best}, and the swarm's centroid are marked. k_χ, bias weight, T, temperature. **Bottom:** Evolution of swarm spread SwS (left) and g_{best} fluctuation (right) during the FLAPS optimization.

Comparison to grid search

We performed comparative grid-search optimizations for the presented protein systems. In swarm-based approaches like FLAPS, each generation depends on the particles' positions and fitnesses attained in the previous generation (see velocity update in Algorithm 1). This recursive dependency limits the intrinsic parallelizability of such algorithms. In contrast, a grid search is perfectly parallel, which hinders a direct speed-up comparison. To ensure comparability of the two approaches, we considered $n_{k_\chi} \times n_T = 15 \times 10 = 150$ equidistant grid points within $-11 \leq \log_{10}\left(k_\chi\right) \leq -8$ and $10 \leq T \leq 90$. Such a grid equates to the sample size and computational demands of the presented FLAPS optimizations. Grid search does not include prior information about the optimization task. Hence, it lacks an intrinsic quality measure to rank different parameter combinations in terms of their performance. The best solutions must be manually identified from the set of evaluated grid points using independently defined quality criteria. PSO implements an intrinsic quality measure in form of the objective function that is key in selecting the parameter combinations to be tested during the optimization process.

We employed the flexible objective function in equation (2) to evaluate the simulated ensemble of protein structures at each grid point. As anticipated, the grid searches yielded several acceptable or even equally functional MD parameter combinations as FLAPS (see Table S5). As proteins are intrinsically dynamic, we are interested in conformational ensembles rather than in single static structures. A PSO search guided by the collective experience of all particles cooperating in a swarm will always yield an overall greater proportion of usable parameter combinations, i.e., meaningful simulations and thus molecular structures, than an exhaustive grid search on a predefined set of parameter combinations. While PSO tends to remember and return to promising regions in the search space, a grid search always takes the risk of evaluating a significant number of ill-suited parameter combinations that would be dismissed on the basis of the swarm's experience.

To illustrate the advantages of a swarm-based approach over classical grid search, we considered the distributions of the simulations' median global distance test, $\mathrm{GDT}^{\mathrm{med}}$, in each grid-search and FLAPS optimization. The distributions are shown in Supplementary Figs. S5 to S12. For all systems, we find the FLAPS distribution to be shifted to higher GDTs compared to the grid search, indicating an overall greater proportion of accurate structural ensembles in FLAPS. In addition, we calculated the fraction of simulations with a median GDT equal to or greater than $\mathrm{GDT}_{75}^{\mathrm{max}} = 0.75 \max\left(\mathrm{GDT}^{\mathrm{med}}\right)$ for each optimization run, which was consistently larger in FLAPS.

System	**LAO ha**	**ADK oc**	**LAO ah**	**ADK co**
Best simulation in terms of OF				
f^{min}	-2.60	-2.14	-2.24	-2.74
k_χ	4.394 e-11	1.390 e-09	8.483 e-10	7.197 e-11
T	10.00	10.00	10.00	10.00
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.2	3.0	2.2	4.3
$\mathrm{GDT}^{\mathrm{med}}$	69.70	62.62	66.49	45.44
Best simulation in terms of $\mathrm{RMSD}^{\mathrm{med}}$				
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.1	2.6	2.0	3.5
$f\left(\mathrm{RMSD}^{\mathrm{med}}\right)$	-2.24	-1.66	-2.11	-1.36
k_χ	3.162 e-10	3.728 e-09	1.390 e-09	1.390 e-09
T	10.00	10.00	10.00	10.00
$\mathrm{GDT}^{\mathrm{med}}$	70.91	63.20	69.33	50.00
Best simulation in terms of $\mathrm{GDT}^{\mathrm{med}}$				
$\mathrm{GDT}^{\mathrm{med}}$	70.91	63.78	69.33	50.47
$f\left(\mathrm{GDT}^{\mathrm{med}}\right)$	-2.24	-2.03	-2.11	-1.97
k_χ	3.162 e-10	2.276 e-09	1.390 e-09	4.394 e-11
T	10.00	10.00	10.00	10.00
$\mathrm{RMSD}^{\mathrm{med}}$ (Å)	2.1	2.7	2.0	4.9

Table S5. Grid-search optimization results. LAO ha, holo-to-apo transition of lysine-, arginine-, ornithine-binding protein, ADK oc, open-to-closed transition of adenylate kinase, LAO ah, apo-to-holo transition of lysine-, arginine-, ornithine-binding protein, ADK co, closed-to-open transition of adenylate kinase, OF, objective function f, k_χ, bias weight, T, temperature, $\mathrm{RMSD}^{\mathrm{med}}$, median root-mean-square deviation, $\mathrm{GDT}^{\mathrm{med}}$, median global distance test (total score).

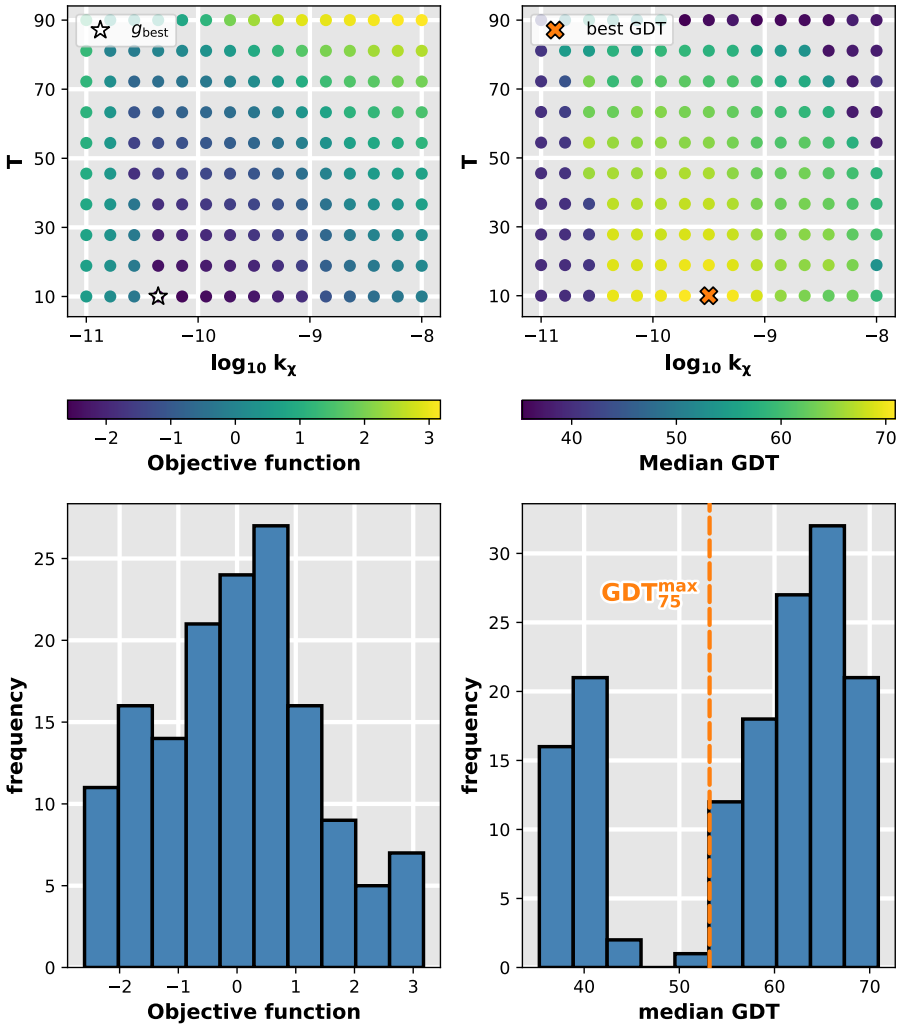

Figure S5. Grid-search optimization for SAXS-guided holo-to-apo transition of lysine-, arginine-, ornithine-binding protein. Top: Final topologies of objective function (OF, left) and median global distance test (GDT, right). The global best position according to the OF, g_{best}, and the best position according to the median GDT are marked. **Bottom:** Frequency distributions of OF (left) and median GDT (right). 73 % of all simulations had a median GDT greater than $GDT_{75}^{max} = 0.75 \max\left(GDT^{med}\right)$.

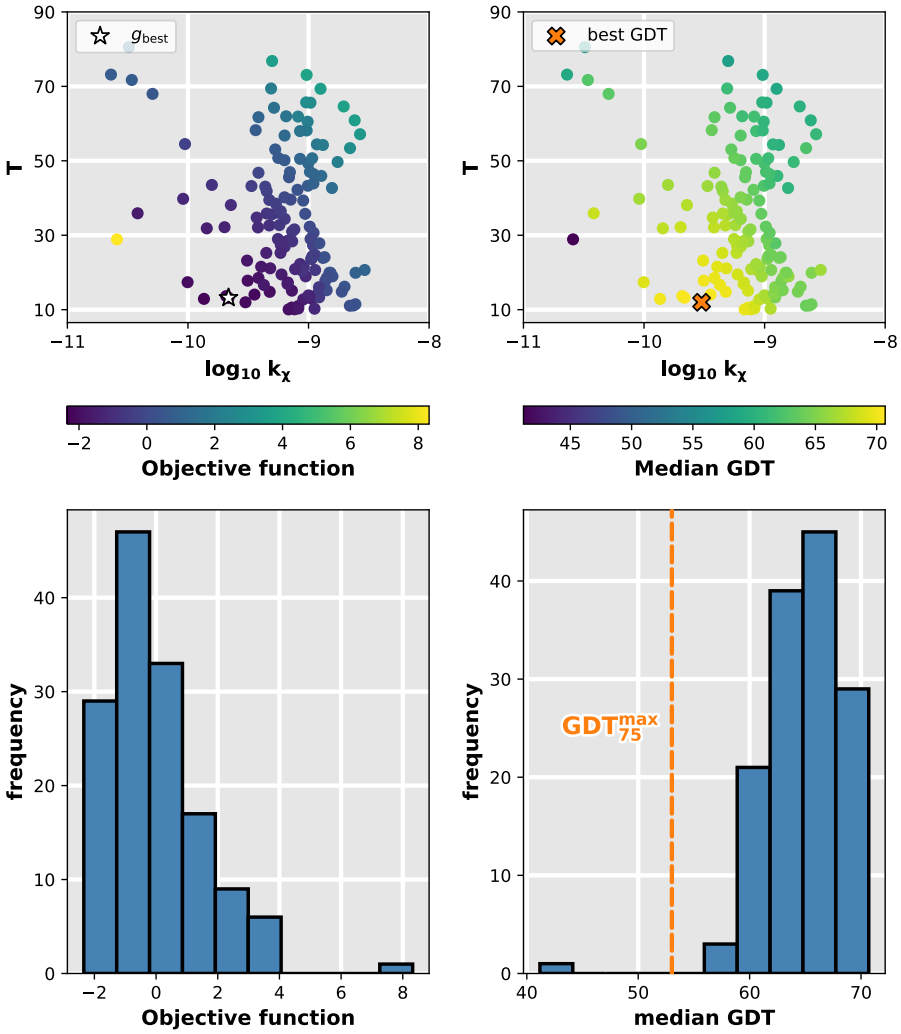

Figure S6. `FLAPS` optimization for SAXS-guided holo-to-apo transition of lysine-, arginine-, ornithine-binding protein. Top: Final topologies of objective function (OF, left) and median global distance test (GDT, right). The global best position according to the OF, g_{best}, and the best position according to the median GDT are marked. **Bottom:** Frequency distributions of OF (left) and median GDT (right). 99 % of all simulations had a median GDT greater than $GDT_{75}^{max} = 0.75 \max\left(GDT^{med}\right)$.

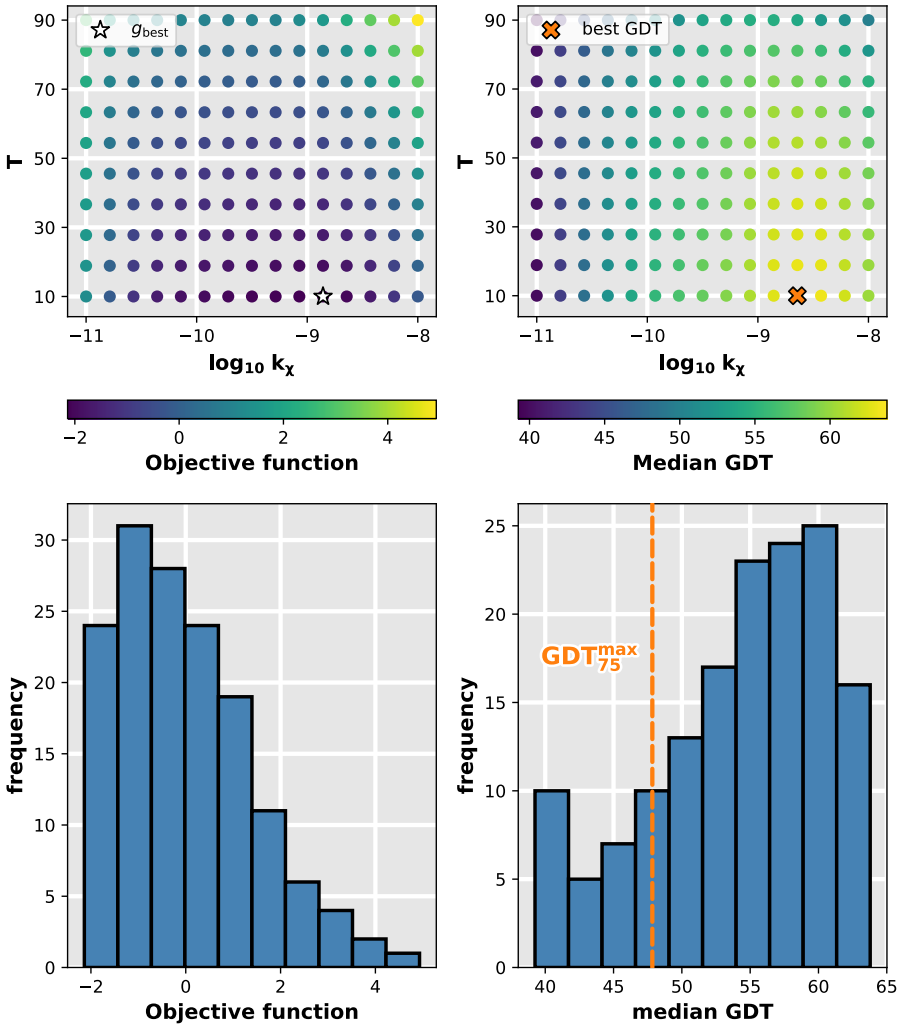

Figure S7. Grid-search optimization for SAXS-guided open-to-closed transition of adenylate kinase. Top: Final topologies of objective function (OF, left) and median global distance test (GDT, right). The global best position according to the OF, g_{best}, and the best position according to the median GDT are marked. **Bottom:** Frequency distributions of OF (left) and median GDT (right). 83 % of all simulations had a median GDT greater than $\text{GDT}_{75}^{\text{max}} = 0.75 \max\left(\text{GDT}^{\text{med}}\right)$.

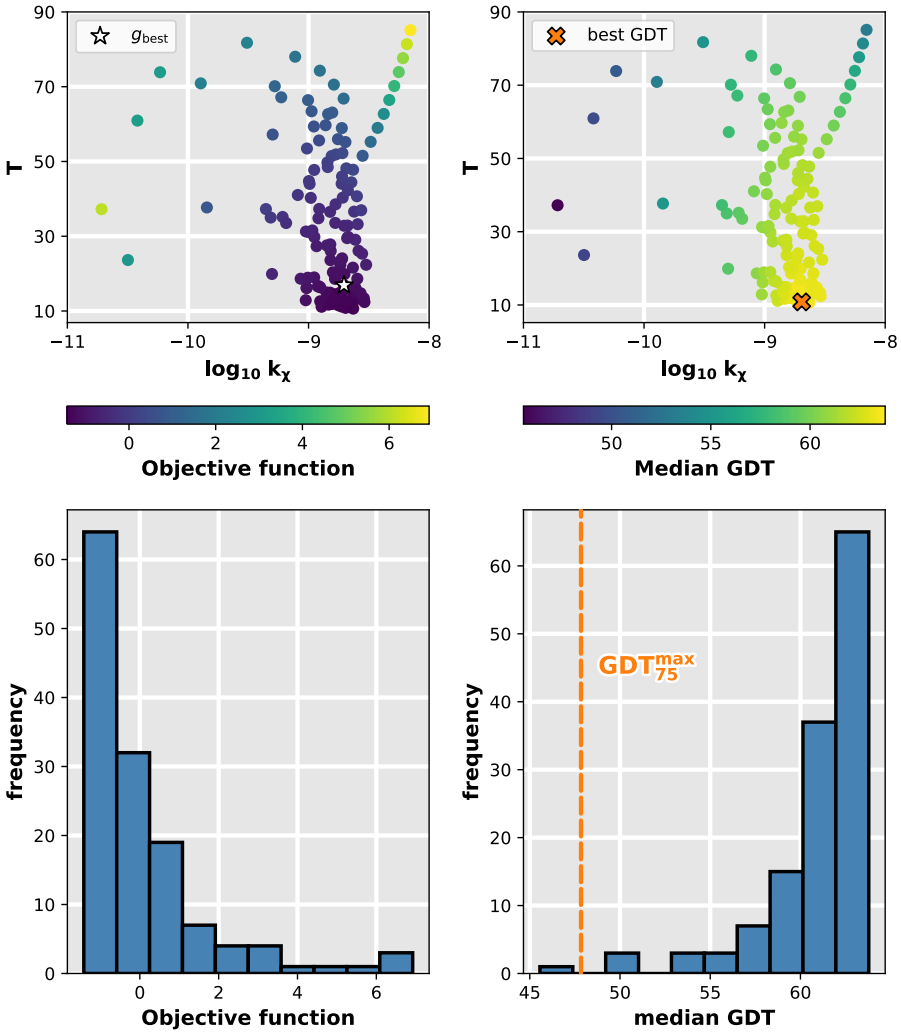

Figure S8. **FLAPS optimization for SAXS-guided open-to-closed transition of adenylate kinase. Top:** Final topologies of objective function (OF, left) and median global distance test (GDT, right). The global best position according to the OF, g_{best}, and the best position according to the median GDT are marked. **Bottom:** Frequency distributions of OF (left) and median GDT (right). 99 % of all simulations had a median GDT greater than $GDT_{75}^{max} = 0.75 \max\left(GDT^{med}\right)$.

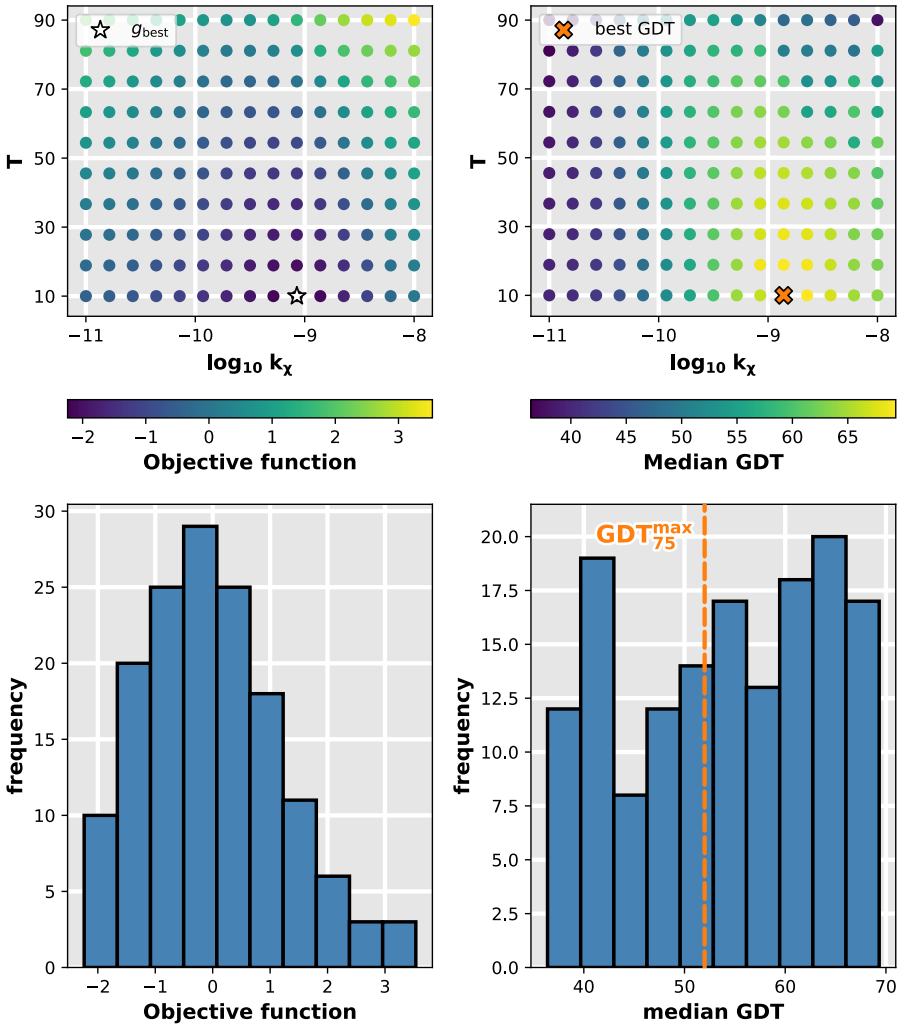

Figure S9. Grid-search optimization for SAXS-guided apo-to-holo transition of lysine-, arginine-, ornithine-binding protein. **Top:** Final topologies of objective function (OF, left) and median global distance test (GDT, right). The global best position according to the OF, g_{best}, and the best position according to the median GDT are marked. **Bottom:** Frequency distributions of OF (left) and median GDT (right). 59 % of all simulations had a median GDT greater than $GDT_{75}^{max} = 0.75 \max\left(GDT^{med}\right)$.

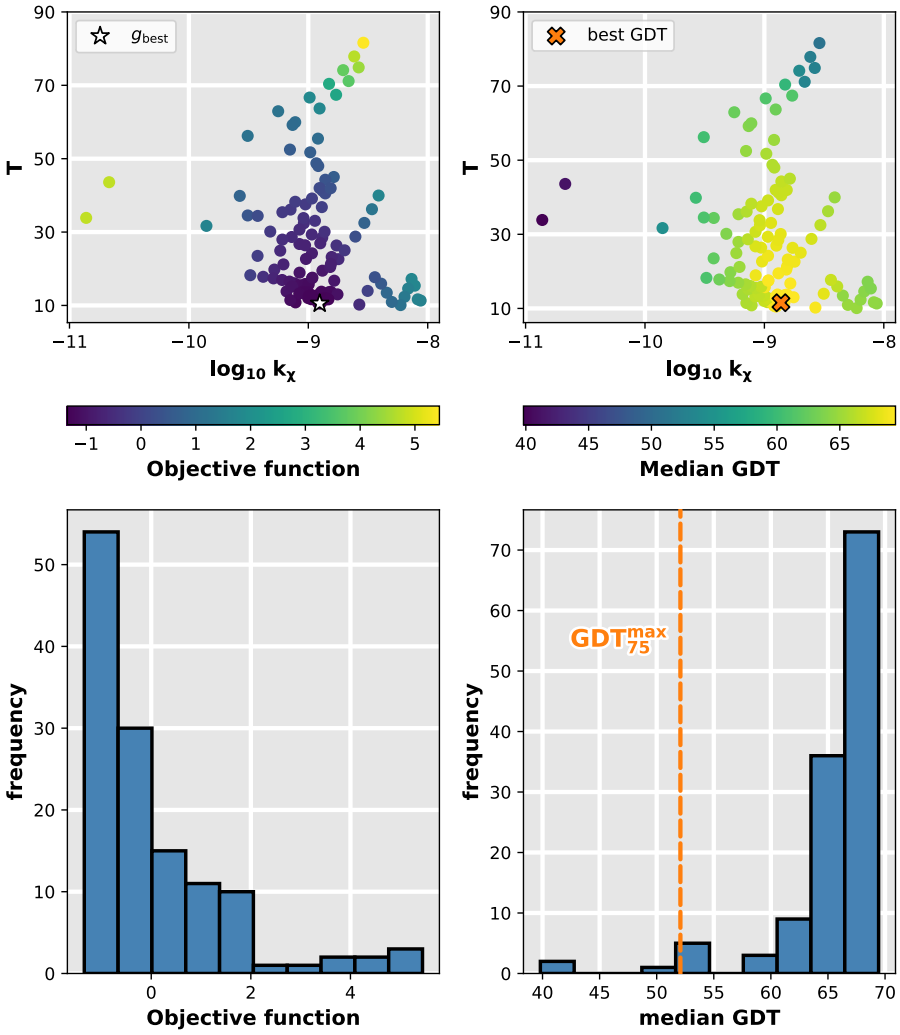

Figure S10. **FLAPS optimization for SAXS-guided apo-to-holo transition of lysine-, arginine-, ornithine-binding protein.** **Top:** Final topologies of objective function (OF, left) and median global distance test (GDT, right). The global best position according to the OF, g_{best}, and the best position according to the median GDT are marked. **Bottom:** Frequency distributions of OF (left) and median GDT (right). 98 % of all simulations had a median GDT greater than $GDT_{75}^{max} = 0.75 \max\left(GDT^{med}\right)$.

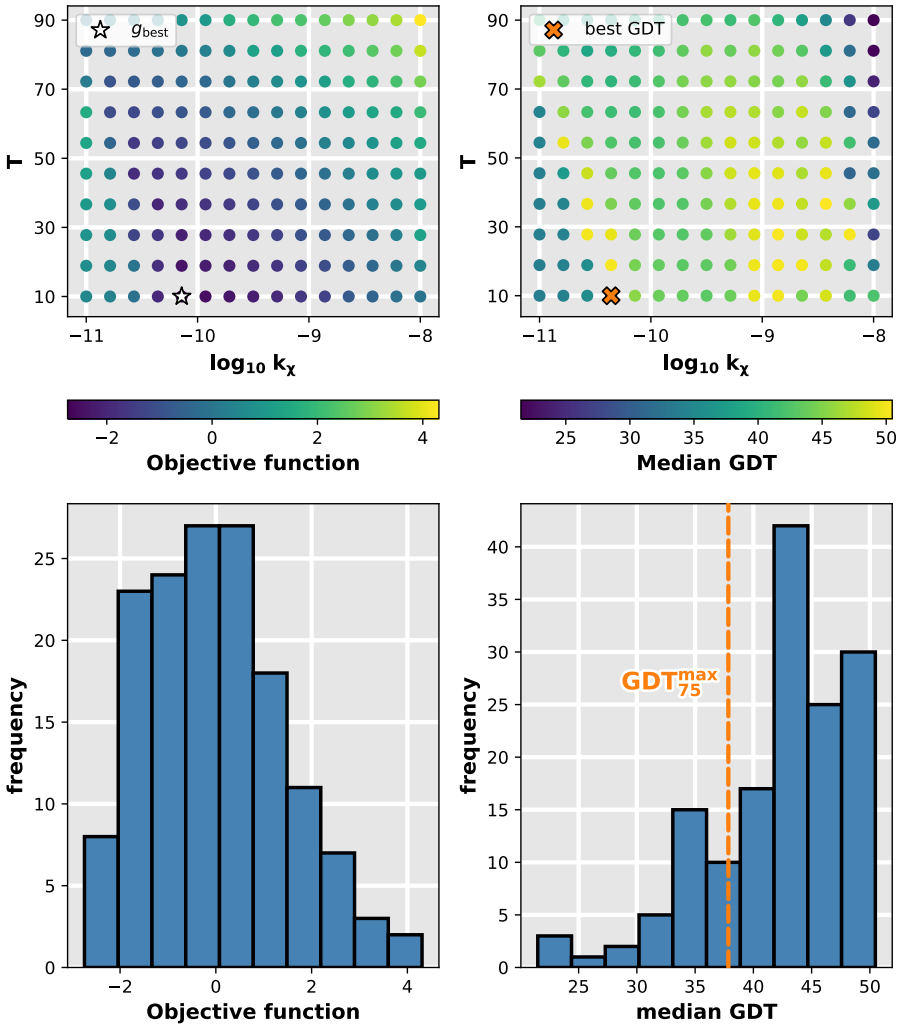

Figure S11. Grid-search optimization for SAXS-guided closed-to-open transition of adenylate kinase. Top: Final topologies of objective function (OF, left) and median global distance test (GDT, right). The global best position according to the OF, g_{best}, and the best position according to the median GDT are marked. **Bottom:** Frequency distributions of OF (left) and median GDT (right). 78 % of all simulations had a median GDT greater than $GDT_{75}^{max} = 0.75 \max\left(GDT^{med}\right)$.

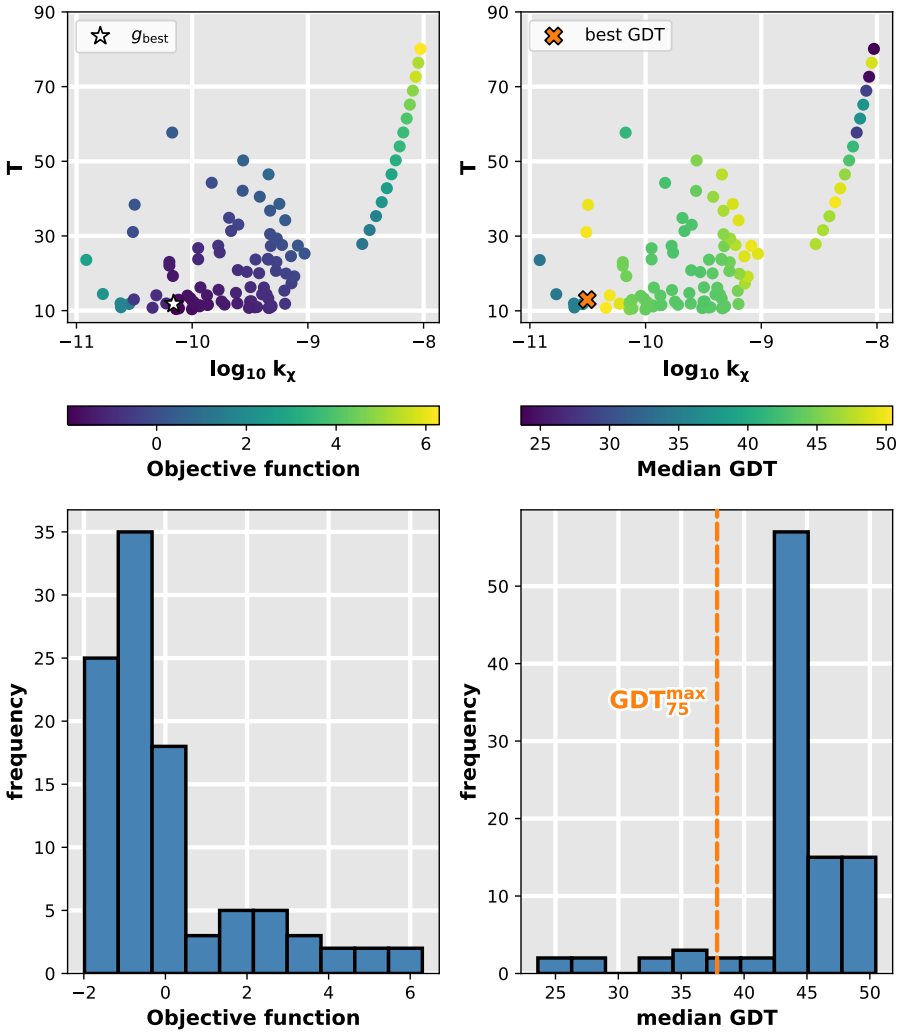

Figure S12. FLAPS optimization for SAXS-guided closed-to-open transition of adenylate kinase. Top: Final topologies of objective function (OF, left) and median global distance test (GDT, right). The global best position according to the OF, g_{best}, and the best position according to the median GDT are marked. **Bottom:** Frequency distributions of OF (left) and median GDT (right). 89 % of all simulations had a median GDT greater than $GDT_{75}^{max} = 0.75 \max\left(GDT^{med}\right)$.

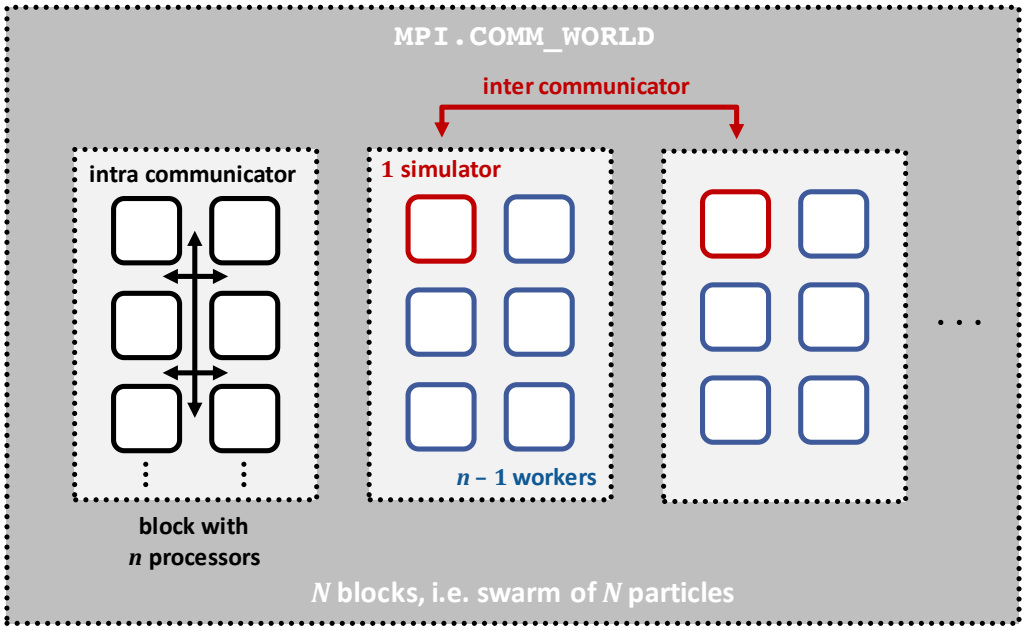

Figure S13. Simulator-worker parallelization scheme used for `FLAPS` in `Hyppopy`. The available compute resources comprising a given number of processors are divided into blocks, each of which corresponds to one particle in the swarm. Within one block, the simulation itself runs on a single core, while the other cores process the generated frames in the trajectory on the fly.