Supporting Information for:

A Molecular Evolution Algorithm for Ligand Design in DOCK

Lauren E. Prentis,[§,1] Courtney D. Singleton,[§,2] John D. Bickel,[3] William J. Allen,[4] Robert C. Rizzo*[,4,5,6]

[1] *Department of Biochemistry & Cell Biology, Stony Brook University, Stony Brook, New York 11794, USA.*

[2] *Department of Molecular Pharmacology, Stony Brook University, Stony Brook, New York, 11794, USA.*

[3] *Department of Chemistry, Stony Brook University, Stony Brook, New York, 11794, USA.*

[4] *Department of Applied Mathematics & Statistics, Stony Brook University, Stony Brook, New York 11794, USA.*

[5] *Institute of Chemical Biology & Drug Discovery, Stony Brook University, Stony Brook, New York 11794, USA.*

[6] *Laufer Center for Physical & Quantitative Biology, Stony Brook University, Stony Brook, New York 11794, USA.*

* Corresponding author e-mail: rizzorc@gmail.com, phone: 631-632-9340, fax: 631-632-8490.

[§] These authors contributed equally to this work.

Contract/grant sponsor: National Institutes of Health; Contract/grant numbers: T32GM136572 (to L.E.P and J.D.B), and R35GM126906 (to R.C.R).

Figure S1. Convergence Results Using Multiple Random Seeds. The graph below shows results for four test systems using three different random seeds.

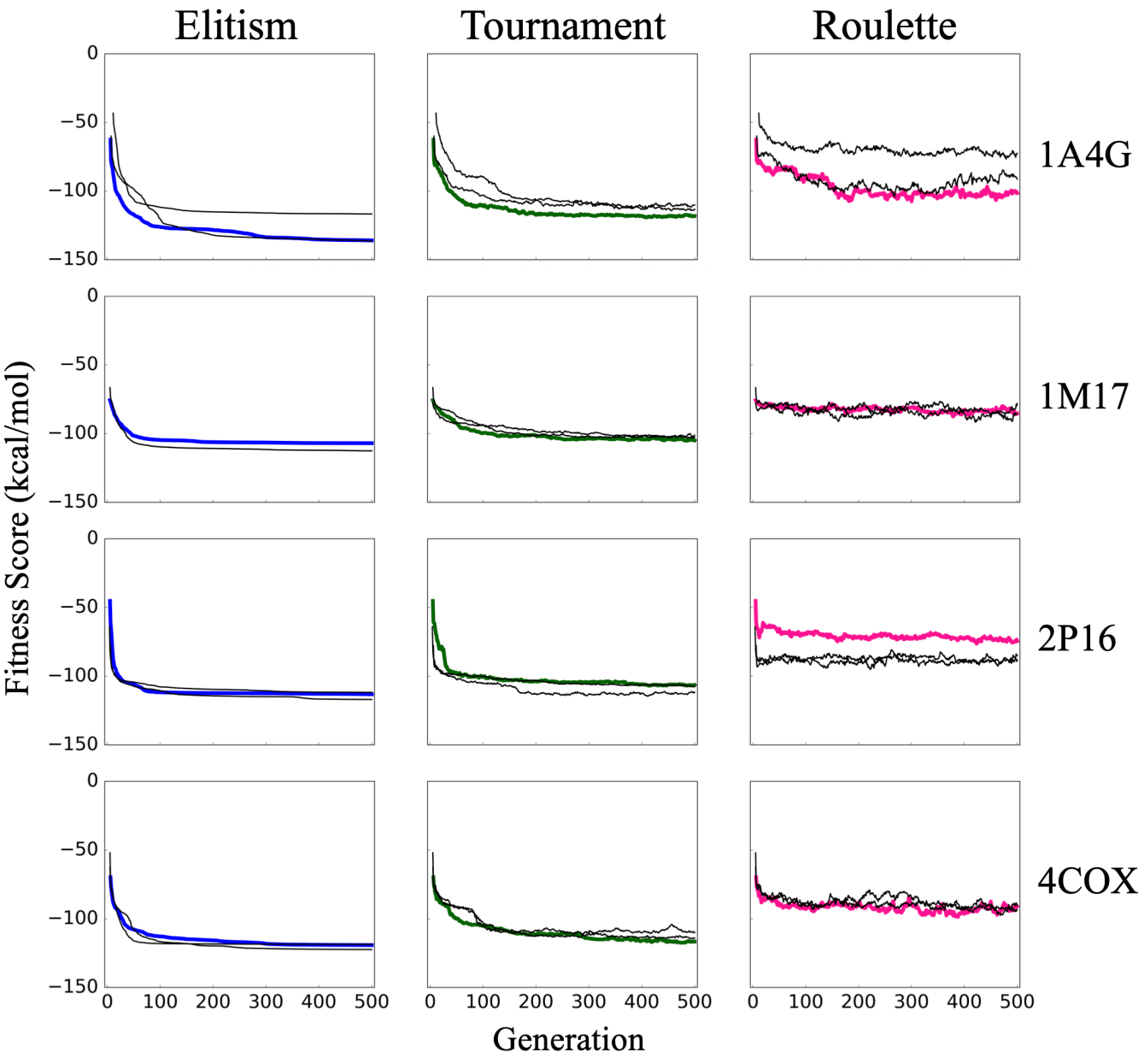

Figure S1. Mean fitness scores with standard deviation, for the top 25 molecules evolved in 1A4G (NA), 1M17 (EGFR), 2P16 (FXA), and 4COX (COX2) using elitism (blue), tournament (green), and roulette (red) selection methods. Dark colored lines (Seed 0) represent the same data as in Figure 8 of the main text. Black lines represent results using two additional random seeds (Seed 2, Seed 3).

Table S1. DOCK6 Genetic Algorithm Input File. The input below is a representative example of a complete parameter file used to execute the genetic algorithm with the elitism selection method described in the main text. Parameters unique to the genetic algorithm are shaded in gray. Interested users should consult the Rizzo lab website for the most up to date DOCK6 tutorials located at https://ringo.ams.stonybrook.edu/index.php/Main_Page.

PARAMETER	VALUE
conformer_search_type	genetic
ga_molecule_file	initial_parent_ensemble.mol2
ga_fraglib_scaffold_file	parameters/fraglib_ga_scaffold.mol2
ga_fraglib_linker_file	parameters/fraglib_linker.mol2
ga_fraglib_sidechain_file	parameters/fraglib_sidechain.mol2
ga_torenv_table	fraglib_torenv.dat
ga_max_generations	100
ga_xover_on	yes
ga_xover_sampling_method_rand	yes
ga_xover_max	150
ga_bond_tolerance	0.5
ga_angle_cutoff	0.14
ga_check_overlap	yes
ga_mutate_addition	yes
ga_mutate_deletion	yes
ga_mutate_substitution	yes
ga_mutate_replacement	yes
ga_use_dn_roulette	no
ga_mutate_parents	yes
ga_pmut_rate	0.4
ga_omut_rate	0.7
ga_max_mut_cycles	10
ga_mut_sampling_method	rand
ga_num_random_picks	10
ga_max_root_size	5
ga_energy_cutoff	100
ga_heur_unmatched_num	5
ga_heur_matched_rmsd	2
ga_constraint_mol_wt	550
ga_constraint_rot_bon	20
ga_constraint_H_accept	10
ga_constraint_H_don	5
ga_constraint_formal_charge	4
ga_ensemble_size	100
ga_selection_method	elitism
ga_elitism_combined	no
ga_elitism_option	max

ga_niching	no
ga_selection_extinction	no
ga_max_num_gen_with_no_crossover	500
ga_name_identifier	ga
ga_output_prefix	1M17
use_internal_energy	yes
internal_energy_rep_exp	12
internal_energy_cutoff	100
use_database_filter	no
orient_ligand	no
bump_filter	no
score_molecules	yes
contact_score_primary	no
grid_score_primary	no
gist_score_primary	no
multigrid_score_primary	no
dock3.5_score_primary	no
continuous_score_primary	no
footprint_similarity_score_primary	no
pharmacophore_score_primary	no
hbond_score_primary	no
interal_score_primary	no
descriptor_score_primary	yes
descriptor_use_grid_score	yes
descriptor_use_pharmacophore_score	yes
descriptor_use_tanimoto	yes
descriptor_use_hungarian	yes
descriptor_use_volume_overlap	yes
descriptor_use_gist	no
descriptor_use_dock3.5	no
descriptor_grid_score_rep_rad_scale	1
descriptor_grid_score_vdw_scale	1
descriptor_grid_score_es_scale	1
descriptor_grid_score_grid_prefix	grid.rec
descriptor_fms_score_use_ref_mol2	yes
descriptor_fms_score_ref_mol2_filename	reference_lig.mol2
descriptor_fms_score_write_reference_pharmacophore_mol2	no
descriptor_fms_score_write_reference_pharmacophore_txt	no
descriptor_fms_score_write_candidate_pharmacophore	no
descriptor_fms_score_write_matched_pharmacophore	no
descriptor_fms_score_compare_type	overlap
descriptor_fms_score_full_match	yes
descriptor_fms_score_match_rate_weight	5
descriptor_fms_score_match_dist_cutoff	1
descriptor_fms_score_match_proj_cutoff	0.7071
descriptor_fms_score_max_score	20
descriptor_fingerprint_ref_filename	reference_lig.mol2
descriptor_hms_score_ref_filename	reference_lig.mol2
descriptor_hms_score_matching_coeff	-5
descriptor_hms_score_rmsd_coeff	1
descriptor_volume_score_reference_mol2_filename	reference_lig.mol2

descriptor_volume_score_overlap_compute_method	analytical
descriptor_weight_grid_score	1
descriptor_weight_pharmacophore_score	0
descriptor_weight_fingerprint_tanimoto	0
descriptor_weight_hms_score	0
descriptor_weight_volume_overlap_score	-15
minimize_ligand	yes
minimize_anchor	yes
minimize_flexible_growth	yes
use_advanced_simplex_parameters	no
simplex_max_cycles	1
simplex_score_converge	0.1
simplex_cycle_converge	1
simplex_trans_step	1
simplex_rot_step	0.1
simplex_tors_step	10
simplex_anchor_max_iterations	500
simplex_grow_max_iterations	250
simplex_grow_tors_premin_iterations	0
simplex_random_seed	4
simplex_restraint_min	yes
simplex_coefficient_restraint	10
atom_model	all
vdw_defn_file	parameters/vdw_de_novo.defn
flex_defn_file	parameters/flex.defn
flex_drive_file	parameters/flex_drive.tbl
chem_defn_file	parameters/chem.defn
pharmacophore_defn_file	parameters/ph4.defn

Note 1: For parameter name "ga_torenv_table" users should augment the standard "fraglib_torenv.dat" supplied with DOCK6 to include any additional torsion entries that the initial parent(s) may contain. Please consult the DOCK6 user's manual for additional information.

Note 2: For parameter name "vdw_defn_file" users should specify the file named "vdw_de_novo.defn" supplied with DOCK6 which includes appropriate VDW parameters for dummy (DU) atoms needed during growth. Please consult the DOCK6 user's manual for additional information.