Supplementary information for:

Macromolecular modeling and design in Rosetta: new methods and frameworks

Rosetta is licensed and distributed through www.rosettacommons.org. Licenses for academic, non-profit and government laboratories are free of charge, there is a license fee for industry users.

The main documentation page can be found at https://www.rosettacommons.org/docs/latest/Home.
Tutorials, demos and protocol captures are documented at https://www.rosettacommons.org/demos/latest/Home.
RosettaScripts documentation is available at https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/RosettaScripts.
PyRosetta tutorials are available at http://www.pyrosetta.org/tutorials.

Method	Use	Developer(s)	Lab developed	Documentation	Protocol capture / demo	Limitations	Competing methods
Protein structure prediction							
fragment picker[1]	picks protein fragments for various modeling tasks	Dominik Gront	Dominik Gront*,**	https://www.rosettacommons.org/docs/latest/application_documentation/utilities/app-fragment-picker	Rosetta/demos/public/fragment_picking		
RosettaCM[2]	comparative modeling from multiple templates	Yifan Song	formerly David Baker	https://www.rosettacommons.org/docs/latest/application_documentation/structure_prediction/RosettaCM	Rosetta/demos/public/homology_modeling_threading_basic Supp to [2]	model quality depends on manual adjustment of sequence alignment, which can take time to do well	MODELLER[3], iTasser[4], HHpred[5]
iterative hybridize[6,7]	recombination of model substructures for *de novo* modeling	Sergey Ovchinnikov, Hahnbeom Park	David Baker, Sergey Ovchinnikov	https://www.rosettacommons.org/docs/latest/IterativeHybridize			
Loop modeling							MODELLER[8]
NGK (next-generation KIC) [9]	*Next Generation Kinematic loop closure*	Amelie Stein	Tanja Kortemme	https://www.rosettacommons.org/docs/latest/application_documentation/structure_prediction/loop_modeling/next-generation-KIC https://www.rosettacommons.org/docs/latest/application_documentation/structure_prediction/loop_modeling/loopmodel-kinematic (for -vicinity_sampling)	Rosetta/demos/tutorials/loop_modeling	performance deteriorates for loops longer than 12 residues, as all loop modeling methods do	Sphinx[10], DiSGro[11]
LoopHashKIC	uses a loop database and KIC for loop design	Xingjie Pan	Tanja Kortemme	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/movers_pages/LoopModelerMover			
Consensus_Loop_Design[12,13]	uses sequence profiles of known loops to design new ones	Enrique Marcos, Tom Linsky	David Baker	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/TaskOperations/taskoperations_pages/ConsensusLoopDesignOperation	Rosetta/main/tests/integration/tests/ConsensusLoopDesign		LOOPY[14], PLOP[15], PS1[16], PS2[17]
GenKIC (generalized KIC) [18]	*Generalized KInematic loop Closure*: robotics based approach to sample non-canonical loop conformations	Vikram Mulligan	Richard Bonneau*	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/composite_protocols/generalized_kic/GeneralizedKIC https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/movers_pages/GeneralizedKICMover https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/composite_protocols/generalized_kic/GeneralizedKICperturber https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/composite_protocols/generalized_kic/GeneralizedKICfilter https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/composite_protocols/generalized_kic/GeneralizedKICselector	https://www.rosettacommons.org/demos/latest/tutorials/GeneralizedKIC/generalized_kinematic_closure_1, https://www.rosettacommons.org/demos/latest/tutorials/GeneralizedKIC/generalized_kinematic_closure_2, https://www.rosettacommons.org/demos/latest/tutorials/GeneralizedKIC/generalized_kinematic_closure_3, https://www.rosettacommons.org/demos/latest/tutorials/GeneralizedKIC/generalized_kinematic_closure_4 Supp to [18,19]	generality requires more setup and can be less efficient than more biased methods; parameter tuning can require some experience	molecular dynamics simulation packages

Method	Use	Developer(s)	Lab developed	Documentation	Protocol capture / demo	Limitations	Competing methods
				https://www.rosettacommons.org/demos/latest/tutorials/GeneralizedKIC/generalized_kinematic_closure_1 https://www.rosettacommons.org/demos/latest/tutorials/GeneralizedKIC/generalized_kinematic_closure_2 https://www.rosettacommons.org/demos/latest/tutorials/GeneralizedKIC/generalized_kinematic_closure_3 https://www.rosettacommons.org/demos/latest/tutorials/GeneralizedKIC/generalized_kinematic_closure_4			
Protein-protein docking							
RosettaDock4.0[20]	protein-protein docking	Nick Marze, Shourya Roy Burman	Jeffrey Gray	https://www.rosettacommons.org/docs/latest/application_documentation/docking/docking-protocol	Rosetta/demos/tutorials/Protein-Protein-Docking https://www.rosettacommons.org/demos/latest/tutorials/Protein-Protein-Docking/Protein-Protein-Docking	global docking requires two orders or magnitude more sampling than local docking – in these cases RosettaDock is slower than FFT-based methods; like for other tools, large conformational changes make docking difficult	InterEVDock[21], ClusPro[22], GalaxyTongDock[23], HADDOCK[24], PPI3D[25]
Rosetta SymDock2[26]	docking of symmetric proteins	Shourya Roy Burman	(formerly Ingemar André), Jeffrey Gray	https://www.rosettacommons.org/docs/latest/SymDockProtocol		SymDock is sensitive to quality of input monomer; as with other algorithms, performance deteriorates for higher-order symmetries	GalaxyTongDock[23], SAM[27], HSYMDOCK[28]
Small molecule ligand docking							
RosettaLigand[29–31]	small molecule docking to proteins	Sam DeLuca, Darwin Fu, Shannon Smith, Rocco Moretti	Jens Meiler	https://www.rosettacommons.org/demos/latest/tutorials/ligand_docking/ligand_docking_tutorial	Rosetta/demos/tutorials/ligand_docking	not as fully a dynamic approach as MD simulations; performs worse in more solvent-exposed pockets or small pockets; requires ligand conformers to be pre-generated; assumes ligand binding site is known	AutoDock Vina[32], GlideScore[33] (Schrodinger), DrugScore[34], Molecular Operating Environment (MOE)[35] from Chemical Computing Group, GOLD[36]
RosettaLigandEnsemble[37]	docks ligand ensembles into proteins	Darwin Fu, Rocco Moretti	Jens Meiler	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/movers_pages/HighResEnsembleMover	Rosetta/demos/protocol_capture/rosettaligand_ensemble	same limitations as RosettaLigand; assumes ligands in ensemble follow similar binding modes	HybridDock[38]
pocket optimization[39,40]	samples binding pocket conformations during docking	David Johnson	John Karanicolas	https://www.rosettacommons.org/docs/latest/application_documentation/utilities/pocket-relax https://www.rosettacommons.org/docs/latest/application_documentation/analysis/pocket-measure	Supp to [39,40]		competitors use small probe molecules (FTMap[41]) or molecular dynamics[42]
DARC[43–45]	*Docking Approach using Ray Casting*: identifies binding pockets	Ragul Gowthaman, Karen Khar	John Karanicolas	https://www.rosettacommons.org/docs/latest/application_documentation/docking/DARC	Rosetta/demos/public/darc Supp to [43–45]		there are numerous approaches (f.ex. DOCKTITE[46] in MOE) for small-molecule docking / virtual screening, but none really comparable to how DARC works
Modeling of antibodies and immune system proteins							
RosettaAntibody[47–50]	antibody homology modeling and docking	Jeliazko Jeliazkov, Nick Marze, Brian Weitzner, Jared Adolf-Bryfogle, Sergey Lyskov, Daisuke Kuroda	Jeffrey Gray	https://www.rosettacommons.org/docs/latest/application_documentation/antibody/antibody-protocol	Rosetta/main/tests/integration/tests/antibody_cc	performance deteriorates for loops longer than 12 residues; unusual CDR loops are hard to model because they are likely not represented in the database	SAbPred[51], PIGS[52], BIOVIA Discovery Studio[53], MOE[54], Schrodinger[55]
AbPredict[56,57]	antibody structure prediction	Christoffer Norn, Gideon Lapidoth	Sarel Fleishman	https://www.rosettacommons.org/docs/latest/application_documentation/structure_prediction/AbPredict2	https://www.rosettacommons.org/demos/latest/tutorials/AbPredict/AbPredict	performance deteriorates for loops longer than 12 residues; unusual CDR loops are hard to model because they are likely not represented in the database	SAbPred[51], PIGS[52], BIOVIA Discovery Studio[53], MOE[54], Schrodinger[55]

Method	Use	Developer(s)	Lab developed	Documentation	Protocol capture / demo	Limitations	Competing methods
RosettaMHC[58]	modeling and docking of antigen / MHC-I / (chaperone or T-cell receptor)	Santrupti Nerli	Nik Sgourakis	https://github.com/snerligit/mhc-pep-threader	https://github.com/snerligit/mhc-pep-threader/tree/master/examples	the binding energies cannot fully discriminate best binders from average binders	bioinformatics-based methods[59]: NetMHCPan[60], MHCFlurry[61] structure-based methods[62]: AutoDock[63], DINC[64]
TCRModel[65]	structure prediction of T-cell receptors	Ragul Gowthaman	Brian Pierce	https://www.rosettacommons.org/docs/latest/application_documentation/structure_prediction/TCRmodel	Rosetta/main/tests/integration/tests/tcrmodel	challenging to model long CDR loops	LYRA[66], ImmuneScape[67]
SnugDock[68]	docking of antibody-antigen complexes	Jeliazko Jeliazkov, Nick Marze, Brian Weitzner	Jeffrey Gray	https://www.rosettacommons.org/docs/latest/application_documentation/antibody/snugdock	Rosetta/main/tests/integration/tests/SnugDock	performance deteriorates for loops longer than 12 residues; cannot predict large backbone motions from unbound to bound state in antigen	usually for FFT methods developers design Ab/Ag specific potentials or paratope/epitope predictions and then combine this with docking methods: ClusPro[69], PatchDock[70], Antibody I-patch[71]
Design of antibodies and immune system proteins							
RAbD[72] (Rosetta AntibodyDesign)	design of antibody-antigen interfaces	Jared Adolf-Bryfogle, Brian Weitzner	Bill Schief, Roland Dunbrack	https://www.rosettacommons.org/docs/latest/application_documentation/antibody/RosettaAntibodyDesign	Rosetta/main/tests/integration/tests/antibody_designer	antigen structure must be known or predicted; well-behaved starting antibody must be given for *de novo* design	OptCDR[73], OptMaven[74]
Epitope removal[75,76]	identifies and removes immunogenic epitopes	Indigo King	David Baker, Cyrus Biotechnology	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/movers_pages/GreedyOptMutationMover https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/xsd/filter_NMerSVMEnergy_type			EpiSweep[77] as part of DisruPPI[78]
AbDesign[79,80]	antibody design	Gideon Lapidoth	Sarel Fleishman	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/SpliceOut https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/SpliceOutAntibody https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/SpliceInAntibody https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/SpliceIn https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/SpliceOutTail https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/SpliceInTail	Rosetta/main/tests/integration/tests/splice_out_H1_H2_longer Rosetta/main/tests/integration/tests/splice_out_H1_H2_same Rosetta/main/tests/integration/tests/splice_out_H1_H2_shorter Rosetta/main/tests/integration/tests/splice_out_H3_longer Rosetta/main/tests/integration/tests/splice_out_H3_same Rosetta/main/tests/integration/tests/splice_out_H3_shorter Rosetta/main/tests/integration/tests/splice_out_L1_L2_longer Rosetta/main/tests/integration/tests/splice_out_L1_L2_same Rosetta/main/tests/integration/tests/splice_out_L1_L2_shorter Rosetta/main/tests/integration/tests/splice_out_L3_longer Rosetta/main/tests/integration/tests/splice_out_L3_same Rosetta/main/tests/integration/tests/splice_out_L3_shorter		deep-learning based methods[81–83]
Protein design							
competitors: review[84], MOE (CCG, Montreal) can be used for manual, not automated, protein design. Supports live energy minimization during design. No appropriate publications							
SEWING[85,86]	*Structure Extension With Native-*	Tim Jacobs, Sharon Guffy, Frank Teets	Brian Kuhlmann	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/composite_protocols/sewing/SEWING		different scores scale differently with size, so top-scoring outputs of a particular run show limited size diversity; restricted to α-	Sibe[87], CCBuilder[88]

Method	Use	Developer(s)	Lab developed	Documentation	Protocol capture / demo	Limitations	Competing methods
	substructure Graphs: recombination of protein substructures to create new designs					helical chimerizable regions; starting point must be single chain; to build from point to point requires additional loop closure step; precludes sampling of backbone motion resulting in high false negative rate	
RosettaRemodel[89]	framework for custom protein remodeling	Possu Huang	Possu Huang*,**	https://www.rosettacommons.org/docs/latest/application_documentation/design/rosettaremodel			
LooDo[90]	LOOp-directed DOmain insertion	Kristin Blacklock	Sagar Khare	https://www.rosettacommons.org/demos/latest/protocol_capture/loodo/README	Rosetta/demos/protocol_capture/loodo		competitors are Remodel, any homology modeling or *de novo* design package; few are specialized for domain insertion in loops
RECON[91]	REstrained CONvergence for multi-state design	Alex Sevy, Marion Sauer	Jens Meiler	https://www.rosettacommons.org/docs/latest/application_documentation/RECON-multistate-design	https://github.com/sevya/msd_analysis_scripts https://github.com/sevya/parallelized_RECON_protocol_capture Supp to [91]	computational requirements; overestimates sequence conservation of the native sequence; under-samples mutations tolerated in related PSI-BLAST profiles during design	no direct competitor; RosettaBackrub[92]; MSD-FASTER[93], dead-end elimination[94], CLASSY[95]
curved β-sheet design[12]	design of curved β-sheets	Enrique Marcos, Benjamin Basanta	David Baker	https://www.rosettacommons.org/docs/latest/application_documentation/design/curvedsheetdesign https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/xsd/filter_StrandCurvatureByLevels_type https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/xsd/filter_HelixBendFilter_type	https://github.com/basantab/DeNovoCurvedSheetDesign	although the fragment assembly should allow construction of arbitrary protein folds, the scripts referenced are limited to NTF2-like proteins; similar setups have been used for generation of *de novo* β-barrel proteins	no known competitors outside of Rosetta
biased forward folding[12]	consistency check for designs	Daniel Silva, Enrique Marcos	David Baker	https://www.rosettacommons.org/docs/latest/Biased-forward-folding https://github.com/emarcos/biased_forward_folding/		not suitable for *ab initio* structure prediction or exhaustive conformational sampling	no known competitors outside of Rosetta
fold_from_loops[96]	designs topology around starting loops	Bruno Correia	Bruno Correia*,**	replaced by FunFolDes below		FunFolDes supersedes fold_from_loops	
FunFolDes[97]	*FUNctional FOLding and DESign:* design topology around a functional motif	Jaume Bonet, Andreas Scheck	Bruno Correia	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/composite_protocols/fold_from_loops/FunFolDes	https://github.com/lpdi-epfl/FunFolDesData https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/composite_protocols/fold_from_loops/RunningFunFolDes	FunFolDes aims to be of use when transferring highly divergent structural motifs, but when the motif and the target region are structurally similar, one would recommend MotifGraft. FunFolDes makes sense to use when the structural divergence is such that it is unavoidable for chain breaks to appear in the final designs.	no known competitors outside of Rosetta; main competitor is MotifGraft[98] in terms of incorporating structural motifs into a topology, even though their goals are different
Protein interface design							
FlexDDG[99]	flexible backbone ΔΔG prediction for interfaces	Kyle Barlow, Shane O'Connor	Tanja Kortemme	https://github.com/Kortemme-Lab/flex_ddG_tutorial	https://github.com/Kortemme-Lab/flex_ddG_tutorial	not designed for ΔΔG prediction for protein stability even though it outperforms *ddg_monomer* which was created for stability prediction	
Coupled Moves[100]	couples flexibility in backbone, sidechains, and ligand for interface design	Noah Ollikainen, Rene M. de Jong	Tanja Kortemme & DSM Biotechnology Center	https://www.rosettacommons.org/docs/latest/coupled-moves	Rosetta/main/tests/integration/tests/coupled_moves		

Method	Use	Developer(s)	Lab developed	Documentation	Protocol capture / demo	Limitations	Competing methods
Parametric design[19,101]	bundle and supercoil design from custom parameters	Vikram Mulligan	Richard Bonneau*	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/movers_pages/MakeBundleMover https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/movers_pages/BundleGridSamplerMover https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/movers_pages/PerturbBundleMover https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Filters/filter_pages/BundleReporterFilter	Supp to [19]	huge parameter space to sample: D^N with D being number of discrete steps and N being number of parameters. For 5 parameters per helix, 10 steps for each parameter and 3 helices, this would be 10^{15} combinations; parameter sampling is only good for topologies that fall within the parameter space	CCBuilder[88]
Peptides and peptidomimetics							
FlexPepDock[102,103]	docking of a flexible peptide into a target protein	Barak Raveh, Nir London, Lior Zimmerman	Ora Schueler-Furman	https://www.rosettacommons.org/docs/latest/application_documentation/docking/flex-pep-dock	Rosetta/demos/public/flex_pep_dock_abinitio Rosetta/demos/public/abinitio_fold_and_dock_of_peptides_using_flexpepdock Rosetta/demos/public/global_dock_ssrA_peptide_against_sspB Rosetta/demos/public/peptide_specificity_using_FlexPepBind Rosetta/demos/public/refinement_of_protein_peptide_complex_using_FlexPepDock Supp to [103]	Both for local refinements, and in particular for *ab initio* docking into a binding site, FlexPepDock is much more CPU expensive than comparable methods. Receptor backbone flexibility has been calibrated for minimization only and therefore conformational changes are of limited extent.	
PIPER-FlexPepDock[104]	flexible peptide docking using FFT-based approach	Nawsad Alam, Alisa Khramushin	Ora Schueler-Furman	https://www.rosettacommons.org/docs/latest/application_documentation/docking/flex-pep-dock		Successful global docking of peptides (starting from peptide sequence and given receptor structure, without any information about the peptide conformation or the peptide binding site on the receptor) is limited to instances where an approximate bound peptide conformation is represented in the fragment set that is rigid body docked in the first step of the docking protocol.	pepATTRACT[105], GalaxyPepDock[106], HADDOCK[107], CABSDock[108], PeptiDock[109], more methods are summarized in this book[110]
PeptiDerive[111]	identifies peptide segment that mediate PPIs	Yuval Sedan, Orly Marcu	Ora Schueler-Furman	https://www.rosettacommons.org/docs/latest/application_documentation/analysis/PeptiDerive	Rosetta/main/tests/integration/tests/peptiderive		LoopFinder[112]
simple_cycpep_predict[18,19,113]	sample peptide conformations through cyclization	Vikram Mulligan	Richard Bonneau*	https://www.rosettacommons.org/docs/latest/structure_prediction/simple_cycpep_predict	Rosetta/main/tests/integration/tests/simple_cycpep_predict and many related demos adjacent, Supp to [18,113]	generality requires more unbiased sampling, therefore slower; limited to 10-12 residue macrocycles, larger only reasonable with additional constraints	molecular dynamics simulation packages
MFPred[114]	mean field approach to predict peptide binding sites for multi-specificity	Aliza Rubenstein	Sagar Khare	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/movers_pages/analysis/GenMeanFieldMover	Rosetta/main/tests/integration/tests/mf_fixbb_sc Rosetta/main/tests/integration/tests/mf_flexbb_sc	dependent on free parameters and optimal parameters depend on the system; no sampling of receptor backbone flexibility	sequence_tolerance[115], PepSpec[116]
RosettaSurface[117–119]	peptide modeling on biomolecular surfaces	Michael Pacella	Jeffrey Gray	https://www.rosettacommons.org/docs/latest/application_documentation/docking/surface-docking	Rosetta/demos/public/surface_docking_cpp Supp to [117]	does not allow 3-body docking	methods reviewed in [120], CHARMM[121], AMBER[122], GRAPPA[123], Interfacial Force Field[124]
Modeling with experimental data							
cryoEM *de novo*[125]	assigns sequence to density *de novo*	Ray Wang	Frank DiMaio, David Baker	https://dimaiolab.ipd.uw.edu/software/	https://dimaiolab.ipd.uw.edu/software/	performance gets a lot worse >600 residues; resolution limit around 4.5Å	Phenix[126], Buccaneer[127]
cryoEM: RosettaES[128]	assigns sequence to density *de novo*, but via enumerative sampling	Brandon Frenz	Frank DiMaio	https://dimaiolab.ipd.uw.edu/software/	https://dimaiolab.ipd.uw.edu/software/		no real competitors (maybe *de novo* tools)

Method	Use	Developer(s)	Lab developed	Documentation	Protocol capture / demo	Limitations	Competing methods
cryoEM: iterative refinement[129]	model refinement into density map	Frank DiMaio, Yifan Song	(formerly David Baker), Frank DiMaio	https://dimaiolab.ipd.uw.edu/software/	https://dimaiolab.ipd.uw.edu/software/	reasonable-quality starting model needs to be available	no real competitors (maybe *de novo* tools)
cryoEM: automated refinement[130]	automated pipeline for refinement into density maps	Ray Wang	Frank DiMaio	https://dimaiolab.ipd.uw.edu/software/	https://dimaiolab.ipd.uw.edu/software/	reasonable-quality starting model needs to be available for all methods (also competing methods) but performance of Rosetta is better than competing methods[130]	Phenix Refine[131], CCP-EM REFMAC[132]
NMR: CS-Rosetta[133]	structure prediction using chemical shifts and RDC/NOE	(formerly Oliver Lange) Santrupti Nerli	Nik Sgourakis	https://csrosetta.chemistry.ucsc.edu/ https://github.com/RosettaCommons/csrosetta3	https://csrosetta.chemistry.ucsc.edu/ https://github.com/RosettaCommons/csrosetta3	depending on the specific needs, setup can be a bit challenging due to lack of good tutorials or protocol captures	competitive methods review:[134]; Xplor-NIH[135], CYANA[136]; similar to autoNOE in CS-Rosetta : ARIA[137], ASDP[138] for automated NOE assignment, NOE interpretation and structure calculation; EC-NMR[139] uses evolutionary couplings
NMR: PCS-Rosetta, GPS-Rosetta[140,141]	structure prediction using PCS and CS data	Christophe Schmitz, Kala Bharath Pilla	Thomas Huber	https://www.rosettacommons.org/demos/latest/protocol_capture/gps_rosetta_pcs_nmr_constraints/README https://github.com/kalabharath/pcs_driven_iterative_resampling	Rosetta/demos/public/gps_rosetta_pcs_nmr_constraints https://github.com/kalabharath/pcs_driven_iterative_resampling/tree/master/sample_run	computational cost for proteins >200 residues can be substantial	for larger proteins, including membrane proteins, DINGO-PCS[142] is used which assembles larger super-secondary structure motifs with PCS
RosettaNMR framework[143]	using RDC/PRE/PCS/NOE/CS for ab initio, protein-protein docking, ligand docking, symmetric assembly	Georg Kuenze, Julia Koehler Leman	Jens Meiler, Richard Bonneau	https://www.rosettacommons.org/docs/latest/application_documentation/RosettaNMR-with-Paramagnetic-Restraints	Rosetta/demos/public/rosettanmr_w_paramagnetic_restraints Supp to [143]	computational cost for proteins >200 residues can be substantial; for systems with many residues, spin-labels, or different metal ions, calculation of tensor parameters can become time-consuming and slow down scoring	competitive methods review:[134]; Xplor-NIH[135], CYANA[136], HADDOCK[144] for docking with paramagnetic restraints or CS,
mass-spec: HRF hydroxyl radical footprinting[145,146]	structure prediction with HRF data	Melanie Aprahamian	Steffen Lindert	https://www.rosettacommons.org/docs/latest/rosetta_basics/scoring/ms_expdata_score_terms	Supp to [145]		
mass-spec: PyTXMS[147]	modeling structures and complexes from MS data	Hamed Khakzad	Lars Malmstroem	https://www.rosettacommons.org/docs/latest/scripting_documentation/PyRosetta/PyTXMS	https://zenodo.org/record/1438111#.XIQUoRNKhBw		no known competitors combine three different MS acquisition data to predict PPI; for analyzing MS/MS samples: StavroX[148] and Mass Spec Studio[149]
DNA and RNA							
RNA applications			Rhiju Das	https://www.rosettacommons.org/docs/latest/application_documentation/rna/rna-applications			
SWA[150,151]	*StepWise Assembly:* RNA modeling one residues at a time	Rhiju Das	Rhiju Das	https://www.rosettacommons.org/docs/latest/application_documentation/stepwise/stepwise_assembly/swa-rna-loop https://www.rosettacommons.org/docs/latest/application_documentation/stepwise/stepwise_assembly/swa-protein-long-loop https://www.rosettacommons.org/docs/latest/application_documentation/stepwise/stepwise_assembly/swa-protein-main	Rosetta/demos/public/swa_rna_loop Rosetta/demos/public/swa_protein_long_loop Rosetta/demos/public/swa_protein_main	superseded by SWM	

Method	Use	Developer(s)	Lab developed	Documentation	Protocol capture / demo	Limitations	Competing methods
SWM[152]	*StepWise Monte-carlo:* SWA using a faster, Monte-Carlo approach	Caleb Geniesse Andrew Watkins	Rhiju Das	https://www.rosettacommons.org/docs/latest/application_documentation/stepwise/stepwise_monte_carlo/stepwise	Rosetta/demos/public/stepwise_enumerative_assembly Rosetta/demos/public/stepwise_monte_carlo_mini_protein Rosetta/demos/public/stepwise_monte_carlo_protein_loop Rosetta/demos/public/stepwise_monte_carlo_rna_loop Rosetta/demos/public/stepwise_monte_carlo_rna_multiloop	generally, greater computational expense than competitive methods	assessment comparison in RNA-puzzles II[153] and III[154]: SimRNA[155], RNAComposer[156], iFoldRNA[157], MC-Sym[158]
FARFAR[159–161]	*Fragment Assembly of RNA with Full-Atom Refinement:* fragment assembly medium resolution structure prediction	Andrew Watkins Kalli Kappel	Rhiju Das	https://www.rosettacommons.org/docs/latest/application_documentation/rna/rna-denovo https://www.rosettacommons.org/docs/latest/application_documentation/rna/rnp-modeling	Rosetta/demos/public/rnp_ddg Rosetta/demos/public/rnp_structure_prediction Rosetta/demos/public/rna_denovo	generally, greater computational expense than competitive methods	assessment comparison in RNA-puzzles II[153] and III[154]: SimRNA[155], RNAComposer[156], iFoldRNA[157], MC-Sym[158]
ERRASER[162,163]	*Enumerative Real-space Refinement ASsisted by Electron density under Rosetta:* refinement into EM density maps	Fang-Chieh Chou	Rhiju Das	https://www.rosettacommons.org/docs/latest/application_documentation/rna/erraser	https://www.rosettacommons.org/demos/latest/public/erraser/README		no known competitors
CS-Rosetta-RNA[164]	RNA modeling with chemical shift NMR data	Parin Sripakdeevong	Rhiju Das	https://www.rosettacommons.org/docs/latest/application_documentation/rna/CS-Rosetta-RNA	Rosetta/demos/public/cs_rosetta_rna		no known competitors
RECCES	*Reweighting of Energy-function Collection with Conformational Ensemble Sampling*	Fang-Chieh Chou	Rhiju Das	https://www.rosettacommons.org/demos/latest/public/recces/README	Rosetta/demos/public/recces		no known competitors
DRRAFTER[165]	*De novo Ribonucleoprotein modeling in Real space through Assembly of Fragments Together with Experimental density in Rosetta: de novo* modeling of protein-RNA complexes into EM densities	Kalli Kappel	Rhiju Das	https://www.rosettacommons.org/docs/latest/application_documentation/rna/drrafter	https://www.rosettacommons.org/demos/latest/public/drrafter/README		no known competitors
Membrane proteins							
limitation: scorefunction models membrane as a symmetric hydrophobic slab of implicit solvent; scorefunctions before franklin2019 do not include pores							
RosettaMP framework[166]: mp_ddg mp_dock mp_relax mp_symdock	membrane protein modeling framework for ΔΔG prediction, protein-protein docking, symmetric docking, and refinement	Julia Koehler Leman, Rebecca Alford	Jeffrey Gray	https://www.rosettacommons.org/docs/latest/application_documentation/Application%20Documentation#Membrane-Proteins	Rosetta/demos/public/mp_ddg Rosetta/demos/public/mp_dock Rosetta/demos/public/mp_relax Rosetta/demos/public/mp_symdock Supp to [166]	all of these methods could be improved in various directions mp_ddg: difficult to get high correlation coefficients mp_dock: only works well for small proteins, needs update to scoring function mp_relax: slow for large proteins	mp_ddg: no known competitors mp_dock: TMDock[167], MemDock[168], PREDDIMER[169] mp_relax: molecular dynamics simulations-based refinement[170]

Method	Use	Developer(s)	Lab developed	Documentation	Protocol capture / demo	Limitations	Competing methods
						mp_symdock: interface packing between subunits often leads to clashes and high scores	mp_symdock: no known competitors
RosettaMP toolkit[171]: mp_score mp_transform mp_mutate_relax helix_from_sequence	accessory toolkit for scoring, transforming into membrane coordinates, mutations, and helix modeling	Julia Koehler Leman	Jeffrey Gray	https://www.rosettacommons.org/docs/latest/application_documentation/Application%20Documentation#Membrane-Proteins	Rosetta/main/tests/integration/tests/mp_transform Rosetta/demos/protocol_capture/helix_from_sequence Supp to [171]	simple but useful methods that work best in conjunction with major modeling protocols	mp_score: useless on its own, so other methods require their own scoring mp_transform: PPM (OPM)[172], TMDET (PDBTM)[173] mp_mutate_relax: no known competitors helix_from_sequence: MODELLER[174]
mp_lipid_acc[175]	prediction of lipid accessibility from structure	Julia Koehler Leman	Richard Bonneau	https://www.rosettacommons.org/docs/latest/application_documentation/Application%20Documentation#Membrane-Proteins	Supp to [175]	requires protein to be transformed into membrane coordinate system; mapping 2D hull into 3D can introduce errors	no known competitors
mp_domain_assembly[176]	domain assembly of full-length membrane proteins from known structures or models	Julia Koehler Leman	Richard Bonneau	https://www.rosettacommons.org/docs/latest/application_documentation/Application%20Documentation#Membrane-Proteins	Rosetta/main/tests/integration/tests/mp_domain_assembly Rosetta/main/tests/integration/tests/mp_domain_assembly_FtsQ Supp to [176]	does not yet work on oligomers	no known competitors
RosettaCM for membrane proteins[177]	multi-template comparative modeling for membrane proteins	Brian Bender	Jens Meiler	https://www.rosettacommons.org/docs/latest/application_documentation/Application%20Documentation#Membrane-Proteins	Supp to [177]	model quality depends on manual adjustment of sequence alignment, which can take time to do well	for membrane proteins: MEDELLER[178], MEMOIR[179], mostly for soluble proteins: MODELLER[3], iTasser[4], HHpred[5]
Carbohydrates							
RosettaCarbohydrate framework[180,181]	framework incorporating carbohydrates into modeling: docking, loop modeling, symmetry, refinement into density maps	Jason W. Labonte, Jared Adolf-Bryfogle	Jeffrey Gray, Bill Schief	https://www.rosettacommons.org/docs/latest/application_documentation/carbohydrates/WorkingWithGlycans		hydrogen bonding is not ideal; charged sugars do not score well; lack of PDB standards complicates input; database lacks several sugar modifications and alternative ring sizes; packing is slow for large sugars	GLYCAM[182], CHARMM-GUI[183]
Scorefunction							
REF2015 scorefunction[184,185]	Rosetta Energy Function from 2015: optimized on structures and thermodynamic observables	Hahnbeom Park, Frank DiMaio	Frank DiMaio, David Baker	https://www.rosettacommons.org/docs/latest/rosetta_basics/scoring/score-types		less accurate but orders of magnitude faster than QM methods; poor performance for many-body effects, induced dipoles, metals; less general than molecular mechanics force fields (f.ex. CHARMM[186]) but better for protein and peptide structure prediction	QM methods: Gaussian[187], GAMESS[188] MD force fields: CHARMM[186], Amber[122]
cartesian_ddG[184]	predicting $\Delta\Delta$Gs for stability in cartesian space	Brandon Frenz, Phil Bradley, Yuan Liu	Frank DiMaio, Phil Bradley	https://www.rosettacommons.org/docs/latest/cartesian-ddG			
HBNet[189,190]	samples hydrogen bonding networks for design	Scott Boyken, Jack Maguire	David Baker, Brian Kuhlman	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/movers_pages/HBNetMover	Rosetta/main/tests/integration/tests/hbnet	does not consider water molecules	no known competitors
HBNetEnergy[189]	score term for scoring hydrogen bonding networks	Vikram Mulligan	Richard Bonneau*	https://www.rosettacommons.org/docs/latest/rosetta_basics/scoring/HBNetEnergy	Rosetta/main/tests/integration/tests/hbnet_energy Rosetta/main/tests/integration/tests/hbnet_energy_rosettascripts_linear		

Method	Use	Developer(s)	Lab developed	Documentation	Protocol capture / demo	Limitations	Competing methods
					Rosetta/main/tests/integration/tests/hbnet_energy_symm		
AACompositionEnergy	penalizes deviations from custom amino acid composition during design	Vikram Mulligan	Richard Bonneau*	https://www.rosettacommons.org/docs/latest/rosetta_basics/scoring/AACompositionEnergy	Supp to [113,191]	slows down design by a factor of 2-5 while raising design efficiency; expertise required for effective use	
AARepeatEnergy	penalizes repeat sequences during design	Vikram Mulligan	Richard Bonneau*	https://www.rosettacommons.org/docs/latest/rosetta_basics/scoring/Repeat-stretch-energy		slows down design by a factor of 2-5 while raising design efficiency; expertise required for effective use	
VoidsPenaltyEnergy	penalizes buried voids in a protein during design	Vikram Mulligan	Richard Bonneau*	https://www.rosettacommons.org/docs/latest/rosetta_basics/scoring/VoidsPenaltyEnergy		slows down design by a factor of 2-5 while raising design efficiency; expertise required for effective use	
NetChargeEnergy	penalizes deviations from custom charge during design	Vikram Mulligan	Richard Bonneau*	https://www.rosettacommons.org/docs/latest/rosetta_basics/scoring/NetChargeEnergy		slows down design by a factor of 2-5 while raising design efficiency; expertise required for effective use	
BuriedUnsatPenalty	penalizes buried, unsaturated hydrogen bond donorso or acceptors	Vikram Mulligan	Richard Bonneau*	https://www.rosettacommons.org/docs/latest/rosetta_basics/scoring/BuriedUnsatPenalty		slows down design by a factor of 2-5 while raising design efficiency; expertise required for effective use	
Interfaces							
PyRosetta[192,193]	Python interface to C++ codebase	Sergey Lyskov	Jeffrey Gray	http://www.pyrosetta.org		greater programming knowledge and familiarity with Rosetta C++ codebase	
RosettaScripts[177,194]	XML-scripts interface to C++ codebase	Sarel Fleishman	Sarel Fleishman*,**	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/RosettaScripts	https://www.rosettacommons.org/demos/latest/tutorials/scripting_with_rosettascripts/scripting_with_rosettascripts https://www.rosettacommons.org/demos/latest/tutorials/advanced_scripting_with_rosettascripts/advanced_scripting_with_rosettascripts	not a complete programming language as loops control flow, conditional logic etc. are limited and clunky, yet easy to learn; intended for linear protocols	
InteractiveRosetta[195]	GUI for main PyRosetta applications	Benjamin Walcott	Chris Bystroff	www.github.com/BystroffLab/InteractiveROSETTA (https://github.com/BystroffLab/InteractiveROSETTA/tree/master/InteractiveROSETTA/help)		not highly used; error-prone, especially interface with PyMol; uses outdated version of PyRosetta; not well supported	
Foldit Standalone[196–199]	video game for playing, teaching, and modeling	Seth Cooper, Firas Khatib	Seth Cooper*,**, Firas Khatib*,**	http://fold.it/standalone			ProteinShop[200] (no longer supported), Coot[201] for electron density, UCSF ChimeraX[202] for some applications, maybe PyMol[203]
ROSIE server[204,205]	super-server hosting numerous Rosetta protocols webservers	Sergey Lyskov, Rocco Moretti, Shane O'Connor	Jeffrey Gray	http://rosie.rosettacommons.org/documentation			N/A
Miscellaneous							
Metalloproteins[206,207]	structure prediction or design involving metal ions	Vikram Mulligan	Richard Bonneau*	https://www.rosettacommons.org/docs/latest/rosetta_basics/non_protein_residues/Metals https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/movers_pages/SetupMetalsMover https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/constraint_generators/MetalContactsConstraintGenerator	[207]	Schrodinger, Gaussian, GAMESS provide more accurate modelling of precise metal geometry and energies but QM methods are too slow to allow large scale sequence and conformational sampling that Rosetta allows	Schrodinger[55], Gaussian[187], GAMESS[188]

Method	Use	Developer(s)	Lab developed	Documentation	Protocol capture / demo	Limitations	Competing methods
Waters[208]	modeling explicit water molecules in conjunction with beta_nov16 scoring function	Ryan Pavlovicz	Frank DiMaio	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/Movers/WaterBoxMover			
SimpleMetrics	framework for running various analyses in RosettaScripts with output to score files	Jared Adolf-Bryfogle	Bill Schief	https://www.rosettacommons.org/docs/latest/scripting_documentation/RosettaScripts/SimpleMetrics/SimpleMetrics			
AmbRose	Python tool for interconversion between Amber and Rosetta	Kristin Blacklock, Aliza Rubenstein, Michael Szegedy	Sagar Khare	https://www.rosettacommons.org/docs/latest/AmbRose	Rosetta/tools/AmbRose		
RosettaRC	user-specific configuration for Rosetta setup	Jared Adolf-Bryfogle	William Schief	https://www.rosettacommons.org/docs/latest/rosetta_basics/running-rosetta-with-options#common-options-and-default-user-configuration			N/A

* the main developer(s) in this lab was/were formerly in the lab of David Baker when this application was developed
** the main developer now has his/her own lab

Summary of additional protocols

Predicting protein structures

fragment picker

Since optimizing the fragments for structure prediction can improve model quality, the original fragment picker application was re-implemented as an object-oriented framework that is vastly more flexible and allows incorporation of various types of restraints from secondary structure prediction or experimental data, for instance from NMR chemical shifts[1].

Small molecule ligand docking

RosettaLigand

Recent improvements to the algorithm rely on a low-resolution sampling step via the *TransformMover*, which combines translational and rotational perturbations in a single step, and using scoring grids for energy evaluation[30]. Further, the algorithm allows backbone flexibility, mimicking the induced fit hypothesis[31]. On a benchmark of 43 complexes, this new algorithm demonstrated an enhanced docking success by 10-15% with an effective 30-fold speedup over the original RosettaLigand performance, enabling virtual high-throughput screening (vHTS) of medium-sized ligand libraries in the order of low hundreds of thousands of ligands.

DARC

Alternatively, these pockets can also be used for Rosetta's *Docking Approach using Ray Casting* (DARC[43]) method. DARC uses ray-casting to rapidly position a ligand in the protein surface pocket[43]; by iterating over many candidates, DARC provides a means for very rapid virtual screening. DARC has also been adapted for GPUs[44], and the newer implementation[45] includes features that provide improved performance in virtual screening benchmarks.

Protein design

LooDo

Multifunctional proteins such as biosensors, bioswitches and tunable affinity clamps can be designed via loop-directed domain insertion (LooDo). Loodo allows inserting a domain into another by two flanking linker regions[90], which are sampled *via* fragment insertion to determine relative positioning of the domains. This is followed by generalized kinematic loop closure[18] (GenKIC, see below) and enzyme design to optimize the interface.

curved beta-sheet design

De novo protein design is somewhat easier for structures consisting of highly regular helices and sheets as their design principles are better understood. In contrast, designing curved and twisted β-sheets requires a deeper understanding of the structural irregularities that enable them. These principles were implemented in the curved β-sheet design method to design a variety of protein folds with curved sheets (Figure 3A), creating pockets suitable for tailoring ligand-binding and enzymatic active sites[12].

fold_from_loops

This approach has been used for antibodies and for vaccine design[96] using the *fold_from_loops* application, where the functional motif is used as a starting point of an extended structure that is folded following the constraints of a target topology. Iterative refinement is carried out via sequence design and structural relaxation before filtering and human-guided optimization.

Protein interface design

CoupledMoves

Designing ligand-binding interfaces in proteins is challenging due to inaccuracies in the energy function (and implicit solvation), the flexibility of ligands, and the sensitivity of protein-ligand interactions to even subtle conformational changes[209]. Flexible backbone design methods that use pre-generated ensembles as a starting point for design[92,115] perform poorly in benchmarks, likely because the ensemble does not accurately describe the unbound-to-bound conformational changes. The *CoupledMoves* protocol couples backbone flexibility with changes in sidechain rotamers or ligand orientation or conformers, and leads to substantial improvements in various benchmarks[210].

Modeling peptides and peptidomimetics

FlexPepDock

FlexPepDock addresses this problem by allowing targeted sampling of the peptide flexibility during its docking into a given binding site, either by refining an approximate peptide conformation (FlexPepDock refinement[102]), or by full *ab initio* sampling of the peptide conformation (FlexPepDock *ab initio*[103]). Peptide docking is especially challenging when the binding site on the receptor is unknown. However, it can be simplified based on the observation that (for peptides built from canonical amino acids) the bound peptide conformation is often included in the fragments generated by the *FragmentPicker*.

PeptiDerive
Many protein-protein interactions (PPI) are mediated by often disordered peptide segments that are responsible for most of the binding energy[112,211–213]. PeptiDerive[111] detects such segments in a PPI complex through a sliding window approach. PeptiDerive was extended to cyclized peptides and is available on the ROSIE[205] server.

MFPred
Multi-specificity is common at protein-peptide interfaces, meaning that the protein can interact with multiple substrates at the same interaction site. This can be exploited for identifying and designing novel substrates. Multi-specificity can be modeled with MFPred[114], which is a rapid, flexible-backbone self-consistent mean field theory-based technique. MFPred can predict experimentally determined peptide specificity profiles for a range of receptors, at equivalent or better prediction accuracy and a 10- to 1000-fold lower computational cost when compared to other methods.

Loop modeling for structure prediction and design
LoopHashKIC + ConsensusLoopDesign
Most Rosetta loop modeling algorithms were primarily developed for structure prediction. However, design constitutes the opposite problem, finding low-energy sequence–structure combinations that satisfy certain design goals. LoopHashKIC[214] addresses this problem and uses the Rosetta LoopHash algorithm[215] to efficiently query a database of loop conformations based on rigid-body transforms between the first and last loop residues. LoopHashKIC uses LoopHash to identify a suitable peptide fragment, and then uses KIC to find an exact solution to close the backbone. To improve the local sequence-structure compatibility in *de novo* designed loops, the *ConsensusLoopDesign* task operation accessed through Rosetta Scripts allows a user to restrict the amino acid identities of loops based on sequence profiles of naturally occurring loops with the same region of backbone dihedral angle space (Ramachandran bins)[12,13].

Modeling antibodies and other proteins in the immune system
RosettaMHC
RosettaMHC[216] utilizes homology modeling and energy minimization to predict structural complexes of antigen/MHC-I/(chaperone or TCR) molecules. It is able to predict peptide antigens that bind to all known MHC-I[58] alleles and models peptide/MHC-I structures[58]. RosettaMHC is implemented in PyRosetta.

RosettaTCR
RosettaTCR[65] generates models of TCRs from sequence, via template identification, grafting of loop templates onto framework regions, and minimization and loop refinement. RosettaTCR permits structural insights into TCRs, for example those targeting cancer neoepitopes[217], or to identify feature sets of TCRs from high throughput sequencing. RosettaTCR can be combined with docking to generate models of TCR-peptide-MHC complexes[218] or TCRs in complex with non-peptide antigens bound to MHC-like proteins[219].

Using experimental data to direct modeling
cryoEM *de novo*
A *de novo* method described by Wang et al. applies a model building approach[125] for density maps between 3-5Å that fits fragments into densities and scores their match based on secondary structure, fit with density, loop closure, clashes and consistency between overlapping fragments to assign sequence into densities. While this method requires >70% of the map to be assigned initially, an updated version of this method, the RosettaES[128] enumerative sampling approach, forgoes this requirement.

cryoEM iterative refinement
One method[129] iterates between refinement with Phenix in reciprocal space to physically plausible conformations, and Rosetta in real space, because Rosetta's all-atom scoring function compensates for the lack of high-resolution data, while the density map restrains backbone and side-chain sampling in real space. Refinement can also be seeded from homology models, followed by density-guided rebuilding and refinement of coordinates and B-factors[220].

<u>Limitations of Rosetta protocols that include experimental data into modeling</u>
Overall limitations of incorporating experimental data into Rosetta are somewhat historical as no single framework combines the use of all types of experimental data; the restraint framework is the common denominator for some data types. Utilization of experimental data has been approached by various labs (typically with experimental expertise in these domains) over many years, with developers having different skill levels in software engineering and preference of user interface, so unfortunately no single ideal solution exists. For example, cryoEM is implemented into the C++ restraint framework, which also holds distance restraints, for instance from NMR or EPR. Other types of NMR data were implemented via the CS-Rosetta framework (typically run through a Python interface), except paramagnetic data, which had a separate C++ implementation. The same is true for mass-spec data, one implementation exists in C++, the other as a Python interface

outside of Rosetta. Further limitations are that experimental data is often implemented as separate score terms, slowing down sampling and scoring, and most methods are limited by the size of the biomolecule. However, one of our software's advantages is that its fragment assembly approach is capable to handle sparse data.

Modeling nucleic acids and their interactions with proteins

FARFAR

The *Fragment Assembly of RNA with Full-Atom Refinement* (FARFAR) structure prediction protocol[159,160] also permits working with chemically modified nucleotides, picking fragments for the most chemically similar base available[152]. Homologous fragments can automatically be eliminated from fragment sets to give pseudo-blind prediction results[152]. As another connection of Rosetta's RNA tools to experimental structural biology, ^{1}H NMR data can be used for RNA modeling via the CS-Rosetta-RNA protocol[164].

Scorefunction

AmbRose

Further, a consensus scoring method, which utilizes the semi-orthogonal nature of the Rosetta and Amber energy functions, was developed for model ranking to identify most near-native models[221] from the pool of generated decoys. This approach led to the development of a Python-based tool (AMBRose) for interconversion between Rosetta and Amber models to facilitate consensus scoring.

Additional information about user interfaces

command line

Structure input and output was enhanced by the ability to read and write mmCIF files (via an external library) using the same mechanisms as PDB files, which permits representation of large complexes that are ill-suited for the PDB format (e.g. the ribosome). This comes with the ability to read the Protein Databank's Chemical Component Dictionary, the description of the chemical composition of residues in the officially released PDB structures. Multithreading support has been added, which required a major refactor of its core architecture for thread-safety, allowing shared-memory parallelism. Multithreading is currently available for specific protocols (*simple_cycpep_predict*) with planned expansion to other applications (including the *JobDistributor* jd3).

RosettaScripts

RosettaScripts[194] is a popular scripting interface that uses E**x**tensible **M**arkup **L**anguage (XML) to build fairly complex protocols using core machinery[222]. Comprehensive knowledge of the codebase is unnecessary since most of the underlying modules[177] have been thoroughly documented – documentation is now also generated using XML schema, which validates the RosettaScripts XML files at runtime. RosettaScripts was further extended and generalized to enhance consistency: *ResidueSelectors* enable selection of residues based on specific properties such as chain, amino acid, secondary structure, index, solvent accessible surface area, and others, and can be used in conjunction with *MoveMapFactories*, which control a structure's flexibility during energy minimization. *ResidueSelectors* are also accepted by *TaskOperations* which control side-chain identity and optimization. A more general analysis tool, *SimpleMetrics*, allows custom analyses of models through RosettaScripts and writes the output into the scorefile. The *SimpleMetrics* system is more integrated and robust than previous tools, such as the *InterfaceAnalyzer* or the *FeaturesReporter*.

PyRosetta

PyRosetta[192,193] is a collection of Python bindings to the source code, exposing ~7,400 classes and 88,000 functions. PyRosetta allows custom protocol development that is flexible and fast, but it requires familiarity with the underlying structure of the codebase. Not all of options available in RosettaScripts have corresponding API-level configuration, so in order to take full advantage of those protocols, PyRosetta can now configure objects using RosettaScripts XML. This brings the added advantage of harmonizing the documentation across multiple interfaces.

InteractiveROSETTA

InteractiveROSETTA[195] is a graphical interface for PyRosetta that presents easy-to-use controls for several of the most widely-used protocols alongside a selection system that uses PyMOL as a visualizer. InteractiveROSETTA is capable of interacting with remote servers running a standalone Rosetta install, rendering it easy to incorporate more sophisticated protocols that are not accessible in PyRosetta and/or require significant computational resources.

Foldit Standalone

Foldit Standalone[196,198] is a graphical interface based on the Foldit video game[197,199]. Foldit Standalone provides several interactive structure manipulations, including pulling directly on the structure, rigid body docking, and residue mutation, insertion and deletion. Users can apply hard and soft constraints that guide automated moves such as packing and minimization, and provides real-time scoring updates as the structure changes. Additional features include multiple sequence alignments for template-based modeling, along with electron density-, Ramachandran-, and contact-map

visualizations. Further, scientists and educators can now run their own custom Foldit puzzles for a group of their choosing, a new feature called "Custom Contests"[223].

Usability and debugging

Command line executables accept a -info option, which prints relevant options for the current application in RosettaScripts, and debugging command lines is facilitated by improved error messages. Default, system-wide options (e.g. database paths) can now be specified in a *rosetta.rc* file. Lastly, code development in C++ is now easier with the help of available code templates that create much of the boilerplate code required to extend the software.

References:

1. Gront, D., Kulp, D. W., Vernon, R. M., Strauss, C. E. M. & Baker, D. Generalized Fragment Picking in Rosetta: Design, Protocols and Applications. *PLoS One* **6,** e23294 (2011).
2. Song, Y., Dimaio, F., Wang, R. Y.-R. R., Kim, D. E., Miles, C., Brunette, T., Thompson, J. & Baker, D. High-resolution comparative modeling with RosettaCM. *Structure* **21,** 1735–1742 (2013).
3. Webb, B. & Sali, A. in *Curr. Protoc. Bioinforma.* **54,** 5.6.1-5.6.37 (John Wiley & Sons, Inc., 2016).
4. Yang, J. & Zhang, Y. in *Curr. Protoc. Bioinforma.* **52,** 5.8.1-5.8.15 (John Wiley & Sons, Inc., 2015).
5. Hildebrand, A., Remmert, M., Biegert, A. & Söding, J. Fast and accurate automatic structure prediction with HHpred. *Proteins Struct. Funct. Bioinforma.* **77,** 128–132 (2009).
6. Ovchinnikov, S., Park, H., Varghese, N., Huang, P.-S., Pavlopoulos, G. A., Kim, D. E., Kamisetty, H., Kyrpides, N. C. & Baker, D. Protein structure determination using metagenome sequence data. *Science (80-.).* **355,** 294–298 (2017).
7. Park, H., Ovchinnikov, S., Kim, D. E., Dimaio, F. & Baker, D. Protein homology model refinement by large-scale energy optimization. *Proc. Natl. Acad. Sci. U. S. A.* **115,** 3054–3059 (2018).
8. Fiser, A., Do, R. K. G. & Šali, A. Modeling of loops in protein structures. *Protein Sci.* **9,** 1753–1773 (2000).
9. Stein, A. & Kortemme, T. Improvements to robotics-inspired conformational sampling in rosetta. *PLoS One* **8,** e63090 (2013).
10. Marks, C., Nowak, J., Klostermann, S., Georges, G., Dunbar, J., Shi, J., Kelm, S. & Deane, C. M. Sphinx: merging knowledge-based and ab initio approaches to improve protein loop prediction. *Bioinformatics* **33,** btw823 (2017).
11. Tang, K., Zhang, J. & Liang, J. Fast Protein Loop Sampling and Structure Prediction Using Distance-Guided Sequential Chain-Growth Monte Carlo Method. *PLoS Comput. Biol.* **10,** e1003539 (2014).
12. Marcos, E., Basanta, B., Chidyausiku, T. M., Tang, Y., Oberdorfer, G., Liu, G., Swapna, G. V. T., Guan, R., Silva, D.-A., Dou, J., Pereira, J. H., Xiao, R., Sankaran, B., Zwart, P. H., Montelione, G. T. & Baker, D. Principles for designing proteins with cavities formed by curved β sheets. *Science* **355,** 201–206 (2017).
13. Marcos, E., Chidyausiku, T. M., McShan, A. C., Evangelidis, T., Nerli, S., Carter, L., Nivón, L. G., Davis, A., Oberdorfer, G., Tripsianes, K., Sgourakis, N. G. & Baker, D. De novo design of a non-local β-sheet protein with high stability and accuracy. *Nat. Struct. Mol. Biol.* **25,** 1028–1034 (2018).
14. Xiang, Z., Soto, C. S. & Honig, B. Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7432–7 (2002).
15. Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J. F., Honig, B., Shaw, D. E. & Friesner, R. A. A hierarchical approach to all-atom protein loop prediction. *Proteins Struct. Funct. Bioinforma.* **55,** 351–367 (2004).
16. Park, H. & Seok, C. Refinement of unreliable local regions in template-based protein models. *Proteins* **80,** 1974–86 (2012).
17. Park, H., Lee, G. R., Heo, L. & Seok, C. Protein loop modeling using a new hybrid energy function and its application to modeling in inaccurate structural environments. *PLoS One* **9,** e113811 (2014).
18. Bhardwaj, G., Mulligan, V. K., Bahl, C. D., Gilmore, J. M., Harvey, P. J., Cheneval, O., Buchko, G. W., Pulavarti, S. V. S. R. K., Kaas, Q., Eletsky, A., Huang, P.-S., Johnsen, W. A., Greisen, P. J., Rocklin, G. J., Song, Y., Linsky, T. W., Watkins, A., Rettie, S. A., Xu, X., Carter, L. P., Bonneau, R., Olson, J. M., Coutsias, E., Correnti, C. E., Szyperski, T., Craik, D. J. & Baker, D. Accurate de novo design of hyperstable constrained peptides. *Nature* **538,** 329–335 (2016).
19. Dang, B., Wu, H., Mulligan, V. K., Mravic, M., Wu, Y., Lemmin, T., Ford, A., Silva, D.-A., Baker, D. & DeGrado, W. F. De novo design of covalently constrained mesosize protein scaffolds with unique tertiary structures. *Proc. Natl. Acad. Sci. U. S. A.* **114,** 10852–10857 (2017).
20. Marze, N. A., Roy Burman, S. S., Sheffler, W. & Gray, J. J. Efficient flexible backbone protein–protein docking for challenging targets. *Bioinformatics* **34,** 3461–3469 (2018).
21. Quignot, C., Rey, J., Yu, J., Tufféry, P., Guerois, R. & Andreani, J. InterEvDock2: an expanded server for protein docking using evolutionary and biological information from homology models and multimeric inputs. *Nucleic Acids Res.* **46,** W408–W416 (2018).
22. Padhorny, D., Kazennov, A., Zerbe, B. S., Porter, K. A., Xia, B., Mottarella, S. E., Kholodov, Y., Ritchie, D. W., Vajda, S. & Kozakov, D. Protein-protein docking by fast generalized Fourier transforms on 5D rotational manifolds. *Proc. Natl. Acad. Sci. U. S. A.* **113,** E4286-93 (2016).

23. Park, T., Baek, M., Lee, H. & Seok, C. GalaxyTongDock: Symmetric and asymmetric ab initio protein–protein docking web server with improved energy parameters. *J. Comput. Chem.* **40,** 2413–2417 (2019).
24. van Zundert, G. C. P., Rodrigues, J. P. G. L. M., Trellet, M., Schmitz, C., Kastritis, P. L., Karaca, E., Melquiond, A. S. J., van Dijk, M., de Vries, S. J. & Bonvin, A. M. J. J. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **428,** 720–725 (2016).
25. Dapkūnas, J., Timinskas, A., Olechnovič, K., Margelevičius, M., Dičiūnas, R. & Venclovas, Č. The PPI3D web server for searching, analyzing and modeling protein–protein interactions in the context of 3D structures. *Bioinformatics* **33,** btw756 (2016).
26. Roy Burman, S. S., Yovanno, R. A. & Gray, J. J. Flexible backbone assembly and refinement of symmetrical homomeric complexes. *bioRxiv* 409730 (2018). doi:10.1101/409730
27. Ritchie, D. W., Grudinin, S. & IUCr. Spherical polar Fourier assembly of protein complexes with arbitrary point group symmetry. *J. Appl. Crystallogr.* **49,** 158–167 (2016).
28. Yan, Y., Tao, H. & Huang, S.-Y. HSYMDOCK: a docking web server for predicting the structure of protein homo-oligomers with Cn or Dn symmetry. *Nucleic Acids Res.* **46,** W423–W431 (2018).
29. Meiler, J. & Baker, D. RosettaLigand: protein-small molecule docking with full side-chain flexibility. *Proteins* **65,** 538–48 (2006).
30. DeLuca, S., Khar, K. & Meiler, J. Fully Flexible Docking of Medium Sized Ligand Libraries with RosettaLigand. *PLoS One* **10,** e0132508 (2015).
31. Davis, I. W. & Baker, D. RosettaLigand Docking with Full Ligand and Receptor Flexibility. *J. Mol. Biol.* **385,** 381–392 (2009).
32. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31,** NA-NA (2009).
33. Richard A. Friesner, *, †, Jay L. Banks, ‡, Robert B. Murphy, ‡, Thomas A. Halgren, ‡, Jasna J. Klicic, ‡, ǁ, Daniel T. Mainz, ‡, Matthew P. Repasky, ‡, Eric H. Knoll, †, Mee Shelley, §, Jason K. Perry, §, David E. Shaw, #, Perry Francis, § and & Shenkin‡, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. (2004). doi:10.1021/JM0306430
34. Brady, G. P. & Stouten, P. F. Fast prediction and visualization of protein binding pockets with PASS. *J. Comput. Aided. Mol. Des.* **14,** 383–401 (2000).
35. Vilar, S., Cozza, G. & Moro, S. Medicinal Chemistry and the Molecular Operating Environment (MOE): Application of QSAR and Molecular Docking to Drug Discovery. *Curr. Top. Med. Chem.* **8,** 1555–1572 (2008).
36. Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W. & Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins Struct. Funct. Bioinforma.* **52,** 609–623 (2003).
37. Fu, D. Y. & Meiler, J. RosettaLigandEnsemble: A Small-Molecule Ensemble-Driven Docking Approach. *ACS Omega* **3,** 3655–3664 (2018).
38. Huang, S.-Y., Li, M., Wang, J. & Pan, Y. HybridDock: A Hybrid Protein–Ligand Docking Protocol Integrating Protein- and Ligand-Based Approaches. *J. Chem. Inf. Model.* **56,** 1078–1087 (2016).
39. Johnson, D. K. & Karanicolas, J. Druggable Protein Interaction Sites Are More Predisposed to Surface Pocket Formation than the Rest of the Protein Surface. *PLoS Comput. Biol.* **9,** e1002951 (2013).
40. Johnson, D. K. & Karanicolas, J. Selectivity by Small-Molecule Inhibitors of Protein Interactions Can Be Driven by Protein Surface Fluctuations. *PLOS Comput. Biol.* **11,** e1004081 (2015).
41. Hall, D. R., Kozakov, D. & Vajda, S. Analysis of protein binding sites by computational solvent mapping. *Methods Mol. Biol.* **819,** 13–27 (2012).
42. Lerner, M. G., Meagher, K. L. & Carlson, H. A. Automated clustering of probe molecules from solvent mapping of protein surfaces: new algorithms applied to hot-spot mapping and structure-based drug design. *J. Comput. Aided. Mol. Des.* **22,** 727–36 (2008).
43. Gowthaman, R., Miller, S. A., Rogers, S., Khowsathit, J., Lan, L., Bai, N., Johnson, D. K., Liu, C., Xu, L., Anbanandam, A., Aubé, J., Roy, A. & Karanicolas, J. DARC: Mapping Surface Topography by Ray-Casting for Effective Virtual Screening at Protein Interaction Sites. *J. Med. Chem.* **59,** 4152–4170 (2016).
44. Khar, K. R., Goldschmidt, L. & Karanicolas, J. Fast Docking on Graphics Processing Units via Ray-Casting. *PLoS One* **8,** e70661 (2013).
45. Gowthaman, R., Lyskov, S. & Karanicolas, J. DARC 2.0: Improved Docking and Virtual Screening at Protein Interaction Sites. *PLoS One* **10,** e0131612 (2015).
46. Scholz, C., Knorr, S., Hamacher, K. & Schmidt, B. DOCKTITE—A Highly Versatile Step-by-Step Workflow for Covalent Docking and Virtual Screening in the Molecular Operating Environment. *J. Chem. Inf. Model.* **55,** 398–406 (2015).
47. Sircar, A., Kim, E. T. & Gray, J. J. RosettaAntibody: antibody variable region homology modeling server. *Nucleic Acids Res.* **37,** W474-479 (2009).
48. Weitzner, B. D., Kuroda, D., Marze, N., Xu, J. & Gray, J. J. Blind prediction performance of RosettaAntibody 3.0: Grafting, relaxation, kinematic loop modeling, and full CDR optimization. *Proteins Struct. Funct. Bioinforma.* **82,** 1611–1623 (2014).

49. Weitzner, B. D., Jeliazkov, J. R., Lyskov, S., Marze, N., Kuroda, D., Frick, R., Adolf-Bryfogle, J., Biswas, N., Dunbrack, R. L. & Gray, J. J. Modeling and docking of antibody structures with Rosetta. *Nat. Protoc.* **12,** 401–416 (2017).
50. Sivasubramanian, A., Sircar, A., Chaudhury, S. & Gray, J. J. Toward high-resolution homology modeling of antibody F $_v$ regions and application to antibody-antigen docking. *Proteins Struct. Funct. Bioinforma.* **74,** 497–514 (2009).
51. Dunbar, J., Krawczyk, K., Leem, J., Marks, C., Nowak, J., Regep, C., Georges, G., Kelm, S., Popovic, B. & Deane, C. M. SAbPred: a structure-based antibody prediction server. *Nucleic Acids Res.* **44,** W474–W478 (2016).
52. Marcatili, P., Rosi, A. & Tramontano, A. PIGS: automatic prediction of antibody structures. *Bioinformatics* **24,** 1953–1954 (2008).
53. BIOVIA Discovery Studio — Discngine - Enhancing Life Science Research. at <https://www.discngine.com/discovery-studio>
54. Molecular Operating Environment (MOE) | MOEsaic | PSILO. at <https://www.chemcomp.com/Products.htm>
55. Schrodinger - Biologics Design. at <https://www.schrodinger.com/science-articles/biologics-design>
56. Norn, C. H., Lapidoth, G. & Fleishman, S. J. High-accuracy modeling of antibody structures by a search for minimum-energy recombination of backbone fragments. *Proteins* **85,** 30–38 (2017).
57. Lapidoth, G., Parker, J., Prilusky, J. & Fleishman, S. J. AbPredict 2: a server for accurate and unstrained structure prediction of antibody variable domains. *Bioinformatics* (2018). doi:10.1093/bioinformatics/bty822
58. Toor, J. S., Rao, A. A., McShan, A. C., Yarmarkovich, M., Nerli, S., Yamaguchi, K., Madejska, A. A., Nguyen, S., Tripathi, S., Maris, J. M., Salama, S. R., Haussler, D. & Sgourakis, N. G. A Recurrent Mutation in Anaplastic Lymphoma Kinase with Distinct Neoepitope Conformations. *Front. Immunol.* **9,** 99 (2018).
59. Mei, S., Li, F., Leier, A., Marquez-Lago, T. T., Giam, K., Croft, N. P., Akutsu, T., Smith, A. I., Li, J., Rossjohn, J., Purcell, A. W. & Song, J. A comprehensive review and performance evaluation of bioinformatics tools for HLA class I peptide-binding prediction. *Brief. Bioinform.* (2019). doi:10.1093/bib/bbz051
60. Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B. & Nielsen, M. NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *J. Immunol.* **199,** 3360–3368 (2017).
61. O'Donnell, T. J., Rubinsteyn, A., Bonsack, M., Riemer, A. B., Laserson, U. & Hammerbacher, J. MHCflurry: Open-Source Class I MHC Binding Affinity Prediction. *Cell Syst.* **7,** 129-132.e4 (2018).
62. Antunes, D. A., Abella, J. R., Devaurs, D., Rigo, M. M. & Kavraki, L. E. Structure-based Methods for Binding Mode and Binding Affinity Prediction for Peptide-MHC Complexes. *Curr. Top. Med. Chem.* **18,** 2239–2255 (2018).
63. Rentzsch, R. & Renard, B. Y. Docking small peptides remains a great challenge: an assessment using AutoDock Vina. *Brief. Bioinform.* **16,** 1045–1056 (2015).
64. Antunes, D. A., Moll, M., Devaurs, D., Jackson, K. R., Lizée, G. & Kavraki, L. E. DINC 2.0: A New Protein–Peptide Docking Webserver Using an Incremental Approach. *Cancer Res.* **77,** e55–e57 (2017).
65. Gowthaman, R. & Pierce, B. G. TCRmodel: high resolution modeling of T cell receptors from sequence. *Nucleic Acids Res.* **46,** W396–W401 (2018).
66. Klausen, M. S., Anderson, M. V., Jespersen, M. C., Nielsen, M. & Marcatili, P. LYRA, a webserver for lymphocyte receptor structural modeling. *Nucleic Acids Res.* **43,** W349–W355 (2015).
67. Li, S., Wilamowski, J., Teraguchi, S., van Eerden, F. J., Rozewicki, J., Davila, A., Xu, Z., Katoh, K. & Standley, D. M. in 207–229 (Humana, New York, NY, 2019). doi:10.1007/978-1-4939-9728-2_17
68. Sircar, A. & Gray, J. J. SnugDock: Paratope Structural Optimization during Antibody-Antigen Docking Compensates for Errors in Antibody Homology Models. *PLoS Comput. Biol.* **6,** e1000644 (2010).
69. Brenke, R., Hall, D. R., Chuang, G.-Y., Comeau, S. R., Bohnuud, T., Beglov, D., Schueler-Furman, O., Vajda, S. & Kozakov, D. Application of asymmetric statistical potentials to antibody-protein docking. *Bioinformatics* **28,** 2608–14 (2012).
70. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33,** W363-7 (2005).
71. Krawczyk, K., Baker, T., Shi, J. & Deane, C. M. Antibody i-Patch prediction of the antibody binding site improves rigid local antibody-antigen docking. *Protein Eng. Des. Sel.* **26,** 621–629 (2013).
72. Adolf-Bryfogle, J., Kalyuzhniy, O., Kubitz, M., Weitzner, B. D., Hu, X., Adachi, Y., Schief, W. R. & Dunbrack, R. L. RosettaAntibodyDesign (RAbD): A general framework for computational antibody design. *PLOS Comput. Biol.* **14,** e1006112 (2018).
73. Pantazes, R. J. & Maranas, C. D. OptCDR: a general computational method for the design of antibody complementarity determining regions for targeted epitope binding. *Protein Eng. Des. Sel.* **23,** 849–58 (2010).
74. Li, T., Pantazes, R. J. & Maranas, C. D. OptMAVEn--a new framework for the de novo design of antibody variable region models targeting specific antigen epitopes. *PLoS One* **9,** e105954 (2014).
75. King, C., Garza, E. N., Mazor, R., Linehan, J. L., Pastan, I., Pepper, M. & Baker, D. Removing T-cell epitopes with computational protein design. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 8577–82 (2014).
76. Nivón, L. G., Bjelic, S., King, C. & Baker, D. Automating human intuition for protein design. *Proteins* **82,** 858–66 (2014).

77. Choi, Y., Verma, D., Griswold, K. E. & Bailey-Kellogg, C. in *Methods Mol. Biol.* **1529,** 375–398 (2017).
78. Choi, Y., Furlon, J. M., Amos, R. B., Griswold, K. E. & Bailey-Kellogg, C. DisruPPI: structure-based computational redesign algorithm for protein binding disruption. *Bioinformatics* **34,** i245–i253 (2018).
79. Lapidoth, G. D., Baran, D., Pszolla, G. M., Norn, C., Alon, A., Tyka, M. D. & Fleishman, S. J. AbDesign: An algorithm for combinatorial backbone design guided by natural conformations and sequences. *Proteins* **83,** 1385–406 (2015).
80. Baran, D., Pszolla, M. G., Lapidoth, G. D., Norn, C., Dym, O., Unger, T., Albeck, S., Tyka, M. D. & Fleishman, S. J. Principles for computational design of binding antibodies. *Proc. Natl. Acad. Sci. U. S. A.* **114,** 10900–10905 (2017).
81. Davidsen, K., Olson, B. J., DeWitt, W. S., Feng, J., Harkins, E., Bradley, P. & Matsen, F. A. Deep generative models for T cell receptor protein sequences. *Elife* **8,** (2019).
82. Riesselman, A., Shin, J.-E., Kollasch, A., McMahon, C., Simon, E., Sander, C., Manglik, A., Kruse, A. & Marks, D. Accelerating Protein Design Using Autoregressive Generative Models. *bioRxiv* 757252 (2019). doi:10.1101/757252
83. Liu, G., Zeng, H., Mueller, J., Carter, B., Wang, Z., Schilz, J., Horny, G., Birnbaum, M. E., Ewert, S. & Gifford, D. K. Antibody Complementarity Determining Region Design Using High-Capacity Machine Learning. *bioRxiv* 682880 (2019). doi:10.1101/682880
84. Gainza, P., Nisonoff, H. M. & Donald, B. R. Algorithms for protein design. *Curr. Opin. Struct. Biol.* **39,** 16–26 (2016).
85. Jacobs, T. M., Williams, B., Williams, T., Xu, X., Eletsky, A., Federizon, J. F., Szyperski, T. & Kuhlman, B. Design of structurally distinct proteins using strategies inspired by evolution. *Science* **352,** 687–90 (2016).
86. Guffy, S. L., Teets, F. D., Langlois, M. I. & Kuhlman, B. Protocols for Requirement-Driven Protein Design in the Rosetta Modeling Program. *J. Chem. Inf. Model.* **58,** 895–901 (2018).
87. Cheung, N. J. & Yu, W. Sibe: a computation tool to apply protein sequence statistics to predict folding and design in silico. *BMC Bioinformatics* **20,** 455 (2019).
88. Wood, C. W. & Woolfson, D. N. CCBuilder 2.0: Powerful and accessible coiled-coil modeling. *Protein Sci.* **27,** 103–111 (2018).
89. Huang, P.-S., Ban, Y.-E. A., Richter, F., Andre, I., Vernon, R., Schief, W. R. & Baker, D. RosettaRemodel: A Generalized Framework for Flexible Backbone Protein Design. *PLoS One* **6,** e24109 (2011).
90. Blacklock, K. M., Yang, L., Mulligan, V. K. & Khare, S. D. A computational method for the design of nested proteins by loop-directed domain insertion. *Proteins Struct. Funct. Bioinforma.* **86,** 354–369 (2018).
91. Sevy, A. M., Jacobs, T. M., Crowe, J. E. & Meiler, J. Design of Protein Multi-specificity Using an Independent Sequence Search Reduces the Barrier to Low Energy Sequences. *PLoS Comput. Biol.* **11,** e1004300 (2015).
92. Smith, C. A. & Kortemme, T. Predicting the tolerated sequences for proteins and protein interfaces using RosettaBackrub flexible backbone design. *PLoS One* **6,** e20451 (2011).
93. Allen, B. D. & Mayo, S. L. An efficient algorithm for multistate protein design based on FASTER. *J. Comput. Chem.* **31,** NA-NA (2009).
94. Yanover, C., Fromer, M. & Shifman, J. M. Dead-end elimination for multistate protein design. *J. Comput. Chem.* **28,** 2122–2129 (2007).
95. Negron, C. Multistate Protein Design Using CLEVER and CLASSY. *Methods Enzymol.* **523,** 171–190 (2013).
96. Correia, B. E., Bates, J. T., Loomis, R. J., Baneyx, G., Carrico, C., Jardine, J. G., Rupert, P., Correnti, C., Kalyuzhniy, O., Vittal, V., Connell, M. J., Stevens, E., Schroeter, A., Chen, M., MacPherson, S., Serra, A. M., Adachi, Y., Holmes, M. A., Li, Y., Klevit, R. E., Graham, B. S., Wyatt, R. T., Baker, D., Strong, R. K., Crowe, J. E., Johnson, P. R. & Schief, W. R. Proof of principle for epitope-focused vaccine design. *Nature* **507,** 201–206 (2014).
97. Bonet, J., Wehrle, S., Schriever, K., Yang, C., Billet, A., Sesterhenn, F., Scheck, A., Sverrisson, F., Veselkova, B., Vollers, S., Lourman, R., Villard, M., Rosset, S., Krey, T. & Correia, B. E. Rosetta FunFolDes - A general framework for the computational design of functional proteins. *PLoS Comput. Biol.* **14,** e1006623 (2018).
98. Silva, D.-A., Correia, B. E. & Procko, E. in *Methods Mol. Biol.* **1414,** 285–304 (2016).
99. Barlow, K. A., Ó Conchúir, S., Thompson, S., Suresh, P., Lucas, J. E., Heinonen, M. & Kortemme, T. Flex ddG: Rosetta Ensemble-Based Estimation of Changes in Protein-Protein Binding Affinity upon Mutation. *J. Phys. Chem. B* **122,** 5389–5399 (2018).
100. Ollikainen, N., de Jong, R. M. & Kortemme, T. Coupling Protein Side-Chain and Backbone Flexibility Improves the Re-design of Protein-Ligand Specificity. *PLOS Comput. Biol.* **11,** e1004335 (2015).
101. Lu, P., Min, D., DiMaio, F., Wei, K. Y., Vahey, M. D., Boyken, S. E., Chen, Z., Fallas, J. A., Ueda, G., Sheffler, W., Mulligan, V. K., Xu, W., Bowie, J. U. & Baker, D. Accurate computational design of multipass transmembrane proteins. *Science (80-.).* **359,** 1042–1046 (2018).
102. Raveh, B., London, N. & Schueler-Furman, O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. *Proteins* **78,** 2029–40 (2010).
103. Raveh, B., London, N., Zimmerman, L. & Schueler-Furman, O. Rosetta FlexPepDock ab-initio: Simultaneous Folding, Docking and Refinement of Peptides onto Their Receptors. *PLoS One* **6,** e18934 (2011).
104. Alam, N., Goldstein, O., Xia, B., Porter, K. A., Kozakov, D. & Schueler-Furman, O. High-resolution global peptide-protein docking using fragments-based PIPER-FlexPepDock. *PLoS Comput. Biol.* **13,** e1005905 (2017).
105. Schindler, C. E. M., de Vries, S. J. & Zacharias, M. Fully Blind Peptide-Protein Docking with pepATTRACT. *Structure* **23,** 1507–1515 (2015).

106. Lee, H., Heo, L., Lee, M. S. & Seok, C. GalaxyPepDock: a protein–peptide docking tool based on interaction similarity and energy optimization. *Nucleic Acids Res.* **43,** W431–W435 (2015).
107. Trellet, M., Melquiond, A. S. J. & Bonvin, A. M. J. J. A Unified Conformational Selection and Induced Fit Approach to Protein-Peptide Docking. *PLoS One* **8,** e58769 (2013).
108. Kurcinski, M., Jamroz, M., Blaszczyk, M., Kolinski, A. & Kmiecik, S. CABS-dock web server for the flexible docking of peptides to proteins without prior knowledge of the binding site. *Nucleic Acids Res.* **43,** W419–W424 (2015).
109. Porter, K. A., Xia, B., Beglov, D., Bohnuud, T., Alam, N., Schueler-Furman, O. & Kozakov, D. ClusPro PeptiDock: efficient global docking of peptide recognition motifs using FFT. *Bioinformatics* **33,** 3299–3301 (2017).
110. *Modeling Peptide-Protein Interactions*. **1561,** (Springer New York, 2017).
111. Sedan, Y., Marcu, O., Lyskov, S. & Schueler-Furman, O. Peptiderive server: derive peptide inhibitors from protein-protein interactions. *Nucleic Acids Res.* **44,** W536-41 (2016).
112. Gavenonis, J., Sheneman, B. A., Siegert, T. R., Eshelman, M. R. & Kritzer, J. A. Comprehensive analysis of loops at protein-protein interfaces for macrocycle design. *Nat. Chem. Biol.* **10,** 716–22 (2014).
113. Hosseinzadeh, P., Bhardwaj, G., Mulligan, V. K., Shortridge, M. D., Craven, T. W., Pardo-Avila, F., Rettie, S. A., Kim, D. E., Silva, D.-A., Ibrahim, Y. M., Webb, I. K., Cort, J. R., Adkins, J. N., Varani, G. & Baker, D. Comprehensive computational design of ordered peptide macrocycles. *Science (80-.).* **358,** 1461–1466 (2017).
114. Rubenstein, A. B., Pethe, M. A. & Khare, S. D. MFPred: Rapid and accurate prediction of protein-peptide recognition multispecificity using self-consistent mean field theory. *PLOS Comput. Biol.* **13,** e1005614 (2017).
115. Smith, C. A. & Kortemme, T. Structure-based prediction of the peptide sequence space recognized by natural and synthetic PDZ domains. *J. Mol. Biol.* **402,** 460–74 (2010).
116. King, C. A. & Bradley, P. Structure-based prediction of protein-peptide specificity in Rosetta. *Proteins* **78,** 3437–49 (2010).
117. Pacella, M. S., Koo, D. C. E., Thottungal, R. A. & Gray, J. J. Using the RosettaSurface algorithm to predict protein structure at mineral surfaces. *Methods Enzymol.* **532,** 343–366 (2013).
118. Lubin, J. H., Pacella, M. S. & Gray, J. J. A Parametric Rosetta Energy Function Analysis with LK Peptides on SAM Surfaces. *Langmuir* **34,** 5279–5289 (2018).
119. Pacella, M. S. & Gray, J. J. A Benchmarking Study of Peptide–Biomineral Interactions. *Cryst. Growth Des.* **18,** 607–616 (2018).
120. Dasetty, S., Meza-Morales, P. J., Getman, R. B. & Sarupria, S. Simulations of interfacial processes: recent advances in force field development. *Curr. Opin. Chem. Eng.* **23,** 138–145 (2019).
121. Mackerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25,** 1400–1415 (2004).
122. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25,** 1157–74 (2004).
123. Hughes, Z. E., Tomásio, S. M. & Walsh, T. R. Efficient simulations of the aqueous bio-interface of graphitic nanostructures with a polarisable model. *Nanoscale* **6,** 5438–5448 (2014).
124. Dharmawardhana, C. C., Kanhaiya, K., Lin, T.-J., Garley, A., Knecht, M. R., Zhou, J., Miao, J. & Heinz, H. Reliable computational design of biological-inorganic materials to the large nanometer scale using Interface-FF. *Mol. Simul.* **43,** 1394–1405 (2017).
125. Wang, R. Y.-R., Kudryashev, M., Li, X., Egelman, E. H., Basler, M., Cheng, Y., Baker, D. & DiMaio, F. De novo protein structure determination from near-atomic-resolution cryo-EM maps. *Nat. Methods* **12,** 335–8 (2015).
126. Terwilliger, T. C., Adams, P. D., Afonine, P. V. & Sobolev, O. V. A fully automatic method yielding initial models from high-resolution cryo-electron microscopy maps. *Nat. Methods* **15,** 905–908 (2018).
127. Cowtan, K. The *Buccaneer* software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62,** 1002–1011 (2006).
128. Frenz, B., Walls, A. C., Egelman, E. H., Veesler, D. & DiMaio, F. RosettaES: a sampling strategy enabling automated interpretation of difficult cryo-EM maps. *Nat. Methods* **14,** 797–800 (2017).
129. DiMaio, F., Echols, N., Headd, J. J., Terwilliger, T. C., Adams, P. D. & Baker, D. Improved low-resolution crystallographic refinement with Phenix and Rosetta. *Nat. Methods* **10,** 1102–4 (2013).
130. Wang, R. Y.-R., Song, Y., Barad, B. A., Cheng, Y., Fraser, J. S. & DiMaio, F. Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. *Elife* **5,** (2016).
131. Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A. & Adams, P. D. Real-space refinement in *PHENIX* for cryo-EM and crystallography. *Acta Crystallogr. Sect. D Struct. Biol.* **74,** 531–544 (2018).
132. Nicholls, R. A., Tykac, M., Kovalevskiy, O. & Murshudov, G. N. Current approaches for the fitting and refinement of atomic models into cryo-EM maps using *CCP-EM*. *Acta Crystallogr. Sect. D Struct. Biol.* **74,** 492–505 (2018).
133. Nerli, S. & Sgourakis, N. G. CS-ROSETTA. *Methods Enzymol.* (2018). doi:10.1016/BS.MIE.2018.07.005
134. Nerli, S., McShan, A. C. & Sgourakis, N. G. Chemical shift-based methods in NMR structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **106–107,** 1–25 (2018).

135. Schwieters, C. D., Bermejo, G. A. & Clore, G. M. Xplor-NIH for molecular structure determination from NMR and other data sources. *Protein Sci.* **27,** 26–40 (2018).
136. Güntert, P. & Buchner, L. Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol. NMR* **62,** 453–471 (2015).
137. Linge, J. P., Habeck, M., Rieping, W. & Nilges, M. ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics* **19,** 315–6 (2003).
138. Huang, Y. J., Tejero, R., Powers, R. & Montelione, G. T. A topology-constrained distance network algorithm for protein structure determination from NOESY data. *Proteins* **62,** 587–603 (2006).
139. Tang, Y., Huang, Y. J., Hopf, T. A., Sander, C., Marks, D. S. & Montelione, G. T. Protein structure determination by combining sparse NMR data with evolutionary couplings. *Nat. Methods* **12,** 751–4 (2015).
140. Yagi, H., Pilla, K. B., Maleckis, A., Graham, B., Huber, T. & Otting, G. Three-dimensional protein fold determination from backbone amide pseudocontact shifts generated by lanthanide tags at multiple sites. *Structure* **21,** 883–890 (2013).
141. Schmitz, C., Vernon, R., Otting, G., Baker, D. & Huber, T. Protein structure determination from pseudocontact shifts using ROSETTA. *J. Mol. Biol.* **416,** 668–77 (2012).
142. Pilla, K. B., Otting, G. & Huber, T. Protein Structure Determination by Assembling Super-Secondary Structure Motifs Using Pseudocontact Shifts. *Structure* **25,** 559–568 (2017).
143. Kuenze, G., Bonneau, R., Koehler Leman, J. & Meiler, J. Integrative protein modeling in RosettaNMR from sparse paramagnetic restraints. *bioRxiv* 597872 (2019). doi:10.1101/597872
144. Dominguez, C., Boelens, R. & Bonvin, A. HADDOCK: a protein− protein docking approach based on biochemical or biophysical Information. *J. Am. Chem. Soc.* 1731–1737 (2003). at <http://pubs.acs.org/doi/abs/10.1021/ja026939x%5Cnfile:///Users/djurredejong/Documents/Articles/Papers/2003/Dominguez2003J. Am. Chem. ….pdf%5Cnpapers://df790107-00bb-4d1a-a7c1-3a1abd474f6d/Paper/p1106>
145. Aprahamian, M. L., Chea, E. E., Jones, L. M. & Lindert, S. Rosetta Protein Structure Prediction from Hydroxyl Radical Protein Footprinting Mass Spectrometry Data. *Anal. Chem.* **90,** 7721–7729 (2018).
146. Aprahamian, M. L. & Lindert, S. Utility of Covalent Labeling Mass Spectrometry Data in Protein Structure Prediction with Rosetta. *J. Chem. Theory Comput.* acs.jctc.9b00101 (2019). doi:10.1021/acs.jctc.9b00101
147. Hauri, S., Khakzad, H., Happonen, L., Teleman, J., Malmström, J. & Malmström, L. Rapid determination of quaternary protein structures in complex biological samples. *Nat. Commun.* **10,** 192 (2019).
148. Götze, M., Pettelkau, J., Schaks, S., Bosse, K., Ihling, C. H., Krauth, F., Fritzsche, R., Kühn, U. & Sinz, A. StavroX--a software for analyzing crosslinked products in protein interaction studies. *J. Am. Soc. Mass Spectrom.* **23,** 76–87 (2012).
149. Sarpe, V., Rafiei, A., Hepburn, M., Ostan, N., Schryvers, A. B. & Schriemer, D. C. High Sensitivity Crosslink Detection Coupled With Integrative Structure Modeling in the Mass Spec Studio. *Mol. Cell. Proteomics* **15,** 3071–80 (2016).
150. Sripakdeevong, P., Kladwang, W. & Das, R. An enumerative stepwise ansatz enables atomic-accuracy RNA loop modeling. *Proc. Natl. Acad. Sci.* **108,** 20573–20578 (2011).
151. Das, R. Atomic-accuracy prediction of protein loop structures through an RNA-inspired Ansatz. *PLoS One* **8,** e74830 (2013).
152. Watkins, A. M., Geniesse, C., Kladwang, W., Zakrevsky, P., Jaeger, L. & Das, R. Blind prediction of noncanonical RNA structure at atomic accuracy. *Sci. Adv.* **4,** eaar5316 (2018).
153. Miao, Z., Adamiak, R. W., Blanchet, M.-F., Boniecki, M., Bujnicki, J. M., Chen, S.-J., Cheng, C., Chojnowski, G., Chou, F.-C., Cordero, P., Cruz, J. A., Ferré-D'Amaré, A. R., Das, R., Ding, F., Dokholyan, N. V., Dunin-Horkawicz, S., Kladwang, W., Krokhotin, A., Lach, G., Magnus, M., Major, F., Mann, T. H., Masquida, B., Matelska, D., Meyer, M., Peselis, A., Popenda, M., Purzycka, K. J., Serganov, A., Stasiewicz, J., Szachniuk, M., Tandon, A., Tian, S., Wang, J., Xiao, Y., Xu, X., Zhang, J., Zhao, P., Zok, T. & Westhof, E. *RNA-Puzzles* Round II: assessment of RNA structure prediction programs applied to three large RNA structures. *RNA* **21,** 1066–1084 (2015).
154. Miao, Z., Adamiak, R. W., Antczak, M., Batey, R. T., Becka, A. J., Biesiada, M., Boniecki, M. J., Bujnicki, J. M., Chen, S.-J., Cheng, C. Y., Chou, F.-C., Ferré-D'Amaré, A. R., Das, R., Dawson, W. K., Ding, F., Dokholyan, N. V., Dunin-Horkawicz, S., Geniesse, C., Kappel, K., Kladwang, W., Krokhotin, A., Łach, G. E., Major, F., Mann, T. H., Magnus, M., Pachulska-Wieczorek, K., Patel, D. J., Piccirilli, J. A., Popenda, M., Purzycka, K. J., Ren, A., Rice, G. M., Santalucia, J., Sarzynska, J., Szachniuk, M., Tandon, A., Trausch, J. J., Tian, S., Wang, J., Weeks, K. M., Williams, B., Xiao, Y., Xu, X., Zhang, D., Zok, T. & Westhof, E. RNA-Puzzles Round III: 3D RNA structure prediction of five riboswitches and one ribozyme. *RNA* **23,** 655–672 (2017).
155. Boniecki, M. J., Lach, G., Dawson, W. K., Tomala, K., Lukasz, P., Soltysinski, T., Rother, K. M. & Bujnicki, J. M. SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic Acids Res.* **44,** e63–e63 (2016).
156. Biesiada, M., Purzycka, K. J., Szachniuk, M., Blazewicz, J. & Adamiak, R. W. in *Methods Mol. Biol.* **1490,** 199–215 (2016).
157. Krokhotin, A., Houlihan, K. & Dokholyan, N. V. iFoldRNA v2: folding RNA with constraints: Fig. 1. *Bioinformatics* **31,** 2891–2893 (2015).

158. Parisien, M. & Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **452,** 51–55 (2008).
159. Das, R., Karanicolas, J. & Baker, D. Atomic accuracy in predicting and designing noncanonical RNA structure. *Nat. Methods* **7,** 291–294 (2010).
160. Cheng, C. Y., Chou, F.-C. & Das, R. Modeling Complex RNA Tertiary Folds with Rosetta. *Methods Enzymol.* **553,** 35–64 (2015).
161. Kappel, K. & Das, R. Sampling Native-like Structures of RNA-Protein Complexes through Rosetta Folding and Docking. *Structure* **27,** 140-151.e5 (2019).
162. Chou, F.-C., Sripakdeevong, P., Dibrov, S. M., Hermann, T. & Das, R. Correcting pervasive errors in RNA crystallography through enumerative structure prediction. *Nat. Methods* **10,** 74–76 (2013).
163. Chou, F.-C., Echols, N., Terwilliger, T. C. & Das, R. in 269–282 (Humana Press, New York, NY, 2016). doi:10.1007/978-1-4939-2763-0_17
164. Sripakdeevong, P., Cevec, M., Chang, A. T., Erat, M. C., Ziegeler, M., Zhao, Q., Fox, G. E., Gao, X., Kennedy, S. D., Kierzek, R., Nikonowicz, E. P., Schwalbe, H., Sigel, R. K. O., Turner, D. H. & Das, R. Structure determination of noncanonical RNA motifs guided by 1H NMR chemical shifts. *Nat. Methods* **11,** 413–416 (2014).
165. Kappel, K., Liu, S., Larsen, K. P., Skiniotis, G., Puglisi, E. V., Puglisi, J. D., Zhou, Z. H., Zhao, R. & Das, R. De novo computational RNA modeling into cryo-EM maps of large ribonucleoprotein complexes. *Nat. Methods* **15,** 947–954 (2018).
166. Alford, R. F., Koehler Leman, J., Weitzner, B. D., Duran, A. M., Tilley, D. C., Elazar, A. & Gray, J. J. An Integrated Framework Advancing Membrane Protein Modeling and Design. *PLoS Comput. Biol.* **11,** e1004398 (2015).
167. Lomize, A. L. & Pogozheva, I. D. TMDOCK: An Energy-Based Method for Modeling α-Helical Dimers in Membranes. *J. Mol. Biol.* **429,** 390–398 (2017).
168. Hurwitz, N., Schneidman-Duhovny, D. & Wolfson, H. J. Memdock: an α-helical membrane protein docking algorithm. *Bioinformatics* **32,** 2444–2450 (2016).
169. Polyansky, A. A., Chugunov, A. O., Volynsky, P. E., Krylov, N. A., Nolde, D. E. & Efremov, R. G. PREDDIMER: A web server for prediction of transmembrane helical dimers. *Bioinformatics* **30,** 889–890 (2014).
170. Dutagaci, B., Heo, L. & Feig, M. Structure refinement of membrane proteins via molecular dynamics simulations. *Proteins Struct. Funct. Bioinforma.* **86,** 738–750 (2018).
171. Koehler Leman, J., Mueller, B. K. & Gray, J. J. Expanding the toolkit for membrane protein modeling in Rosetta. *Bioinformatics* **11,** 1–3 (2016).
172. Lomize, A. L., Pogozheva, I. D., Lomize, M. A. & Mosberg, H. I. Positioning of proteins in membranes: A computational approach. *Protein Sci.* **15,** 1318–1333 (2006).
173. Tusnády, G. E., Dosztányi, Z., Simon, I. I., Tusn??dy, G. E., Doszt??nyi, Z., Simon, I. I., Tusn, E., Tusn??dy, G. E., Doszt??nyi, Z. & Simon, I. I. TMDET: Web server for detecting transmembrane regions of proteins by using their 3D coordinates. *Bioinformatics* **21,** 1276–1277 (2005).
174. Fiser, A. & Sali, A. MODELLER: Generation and Refinement of Homology-Based Protein Structure Models. *Methods Enzymol.* **374,** 461–491 (2003).
175. Koehler Leman, J., Lyskov, S. & Bonneau, R. Computing structure-based lipid accessibility of membrane proteins with mp_lipid_acc in RosettaMP. *BMC Bioinformatics* **18,** 115 (2017).
176. Koehler Leman, J. & Bonneau, R. A novel domain assembly routine for creating full-length models of membrane proteins from known domain structures. *Biochemistry* acs.biochem.7b00995 (2017). doi:10.1021/acs.biochem.7b00995
177. Bender, B. J., Cisneros, A., Duran, A. M., Finn, J. A., Fu, D., Lokits, A. D., Mueller, B. K., Sangha, A. K., Sauer, M. F., Sevy, A. M., Sliwoski, G., Sheehan, J. H., Dimaio, F., Meiler, J. & Moretti, R. Protocols for Molecular Modeling with Rosetta3 and RosettaScripts. *Biochemistry* acs.biochem.6b00444 (2016). doi:10.1021/acs.biochem.6b00444
178. Kelm, S., Shi, J. & Deane, C. M. MEDELLER: Homology-based coordinate generation for membrane proteins. *Bioinformatics* **26,** 2833–2840 (2010).
179. Ebejer, J. P., Hill, J. R., Kelm, S., Shi, J. & Deane, C. M. Memoir: template-based structure prediction for membrane proteins. *Nucleic Acids Res.* **41,** 379–383 (2013).
180. Labonte, J. W., Adolf-Bryfogle, J., Schief, W. R. & Gray, J. J. Residue-centric modeling and design of saccharide and glycoconjugate structures. *J. Comput. Chem.* **38,** 276–287 (2017).
181. Frenz, B., Rämisch, S., Borst, A. J., Walls, A. C., Adolf-Bryfogle, J., Schief, W. R., Veesler, D. & DiMaio, F. Automatically Fixing Errors in Glycoprotein Structures with Rosetta. *Structure* **0,** (2018).
182. Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L. & Woods, R. J. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* **29,** 622–655 (2008).
183. Park, S.-J., Lee, J., Qi, Y., Kern, N. R., Lee, H. S., Jo, S., Joung, I., Joo, K., Lee, J. & Im, W. CHARMM-GUI *Glycan Modeler* for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology* **29,** 320–331 (2019).
184. Park, H., Bradley, P., Greisen, P., Liu, Y., Mulligan, V. K., Kim, D. E., Baker, D. & DiMaio, F. Simultaneous Optimization of Biomolecular Energy Functions on Features from Small Molecules and Macromolecules. *J. Chem. Theory Comput.* **12,** 6201–6212 (2016).

185. Alford, R. F., Leaver-Fay, A., Jeliazkov, J. R., O'Meara, M. J., Dimaio, F. P., Park, H., Shapovalov, M. V., Renfrew, P. D., Mulligan, V. K., Kappel, K., Labonte, J. W., Pacella, M. S., Bonneau, R., Bradley, P., Dunbrack, R. L., Das, R., Baker, D., Kuhlman, B., Kortemme, T. & Gray, J. J. The Rosetta all-atom energy function for macromolecular modeling and design. *J. Chem. Theory Comput.* **13,** 1–35 (2017).
186. Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M. & Karplus, M. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **30,** 1545–1614 (2009).
187. Gaussian.com | Expanding the limits of computational chemistry. at <https://gaussian.com/>
188. Alexeev, Y., P. Mazanetz, M., Ichihara, O. & G. Fedorov, D. GAMESS As a Free Quantum-Mechanical Platform for Drug Research. *Curr. Top. Med. Chem.* **12,** 2013–2033 (2012).
189. Boyken, S. E., Chen, Z., Groves, B., Langan, R. A., Oberdorfer, G., Ford, A., Gilmore, J. M., Xu, C., DiMaio, F., Pereira, J. H., Sankaran, B., Seelig, G., Zwart, P. H. & Baker, D. De novo design of protein homo-oligomers with modular hydrogen-bond network-mediated specificity. *Science* **352,** 680–7 (2016).
190. Maguire, J. B., Boyken, S. E., Baker, D. & Kuhlman, B. Rapid Sampling of Hydrogen Bond Networks for Computational Protein Design. *J. Chem. Theory Comput.* **14,** 2751–2760 (2018).
191. Rocklin, G. J., Chidyausiku, T. M., Goreshnik, I., Ford, A., Houliston, S., Lemak, A., Carter, L., Ravichandran, R., Mulligan, V. K., Chevalier, A., Arrowsmith, C. H. & Baker, D. Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science* **357,** 168–175 (2017).
192. Chaudhury, S., Lyskov, S. & Gray, J. J. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* **26,** 689–691 (2010).
193. Gray, J. J., Chaudhury, S., Lyskov, S., and Labonte, J. W. The PyRosetta Interactive Platform for Protein Structure Prediction and Design: A Set of Educational Modules. (2014). at <http://www.amazon.com/PyRosetta-Interactive-Platform-Structure-Prediction/dp/1500968277>
194. Fleishman, S. J., Leaver-Fay, A., Corn, J. E., Strauch, E.-M. M., Khare, S. D., Koga, N., Ashworth, J., Murphy, P., Richter, F., Lemmon, G., Meiler, J. & Baker, D. RosettaScripts: A scripting language interface to the Rosetta Macromolecular modeling suite. *PLoS One* **6,** 1–10 (2011).
195. Schenkelberg, C. D. & Bystroff, C. InteractiveROSETTA: A graphical user interface for the PyRosetta protein modeling suite. *Bioinformatics* (2015). doi:10.1093/bioinformatics/btv492
196. Kleffner, R., Flatten, J., Leaver-Fay, A., Baker, D., Siegel, J. B., Khatib, F. & Cooper, S. Foldit Standalone: a video game-derived protein structure manipulation interface using Rosetta. *Bioinformatics* **33,** 2765–2767 (2017).
197. Khatib, F., Cooper, S., Tyka, M. D., Xu, K., Makedon, I., Popovic, Z., Baker, D. & Players, F. Algorithm discovery by protein folding game players. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 18949–53 (2011).
198. Cooper, S., Sterling, A. L. R., Kleffner, R., Silversmith, W. M. & Siegel, J. B. Repurposing citizen science games as software tools for professional scientists. in *Proc. 13th Int. Conf. Found. Digit. Games - FDG '18* 1–6 (ACM Press, 2018). doi:10.1145/3235765.3235770
199. Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z. & Players, F. Predicting protein structures with a multiplayer online game. *Nature* **466,** 756–760 (2010).
200. Crivelli, S., Kreylos, O., Hamann, B., Max, N. & Bethel, W. ProteinShop: a tool for interactive protein manipulation and steering. *J. Comput. Aided. Mol. Des.* **18,** 271–85 (2004).
201. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 2126–2132 (2004).
202. Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H. & Ferrin, T. E. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27,** 14–25 (2018).
203. Software: The PyMOL Molecular Graphics System, Version 1.8, Schroedinger LLC. at <https://www.schrodinger.com/suites/pymol>
204. Lyskov, S., Chou, F. C., Conch??ir, S. ??, Der, B. S., Drew, K., Kuroda, D., Xu, J., Weitzner, B. D., Renfrew, P. D., Sripakdeevong, P., Borgo, B., Havranek, J. J., Kuhlman, B., Kortemme, T., Bonneau, R., Gray, J. J. & Das, R. Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). *PLoS One* **8,** 5–7 (2013).
205. Moretti, R., Lyskov, S., Das, R., Meiler, J. & Gray, J. J. Web-accessible molecular modeling with Rosetta: The Rosetta Online Server that Includes Everyone (ROSIE). *Protein Sci.* **27,** 259–268 (2018).
206. Mills, J. H., Khare, S. D., Bolduc, J. M., Forouhar, F., Mulligan, V. K., Lew, S., Seetharaman, J., Tong, L., Stoddard, B. L. & Baker, D. Computational Design of an Unnatural Amino Acid Dependent Metalloprotein with Atomic Level Accuracy. *J. Am. Chem. Soc.* **135,** 13393–13399 (2013).
207. Mulligan, V. K. Manuscript in preparation. (2019).
208. Pavlovicz, R. E., Park, H. & DiMaio, F. Efficient consideration of coordinated water molecules improves computational protein-protein and protein-ligand docking. *bioRxiv* 618603 (2019). doi:10.1101/618603
209. Dou, J., Doyle, L., Jr Greisen, P., Schena, A., Park, H., Johnsson, K., Stoddard, B. L. & Baker, D. Sampling and

energy evaluation challenges in ligand binding protein design. *Protein Sci.* **26,** 2426–2437 (2017).
210. Ollikainen, N., de Jong, R. M. & Kortemme, T. Coupling Protein Side-Chain and Backbone Flexibility Improves the Re-design of Protein-Ligand Specificity. *PLoS Comput. Biol.* **11,** e1004335 (2015).
211. London, N., Raveh, B., Movshovitz-Attias, D. & Schueler-Furman, O. Can self-inhibitory peptides be derived from the interfaces of globular protein-protein interactions? *Proteins* **78,** 3140–9 (2010).
212. Watkins, A. M. & Arora, P. S. Anatomy of β-strands at protein-protein interfaces. *ACS Chem. Biol.* **9,** 1747–54 (2014).
213. Jochim, A. L. & Arora, P. S. Systematic analysis of helical protein interfaces reveals targets for synthetic inhibitors. *ACS Chem. Biol.* **5,** 919–23 (2010).
214. Pan, X. & Kortemme, T. Manuscript in preparation. (2019).
215. Tyka, M. D., Jung, K. & Baker, D. Efficient sampling of protein conformational space using fast loop building and batch minimization on highly parallel computers. *J. Comput. Chem.* **33,** 2483–2491 (2012).
216. Nerli, S. & Sgourakis, N. G. Manuscript in preparation. (2019).
217. Tran, E., Robbins, P. F., Lu, Y.-C., Prickett, T. D., Gartner, J. J., Jia, L., Pasetto, A., Zheng, Z., Ray, S., Groh, E. M., Kriley, I. R. & Rosenberg, S. A. T-Cell Transfer Therapy Targeting Mutant KRAS in Cancer. *N. Engl. J. Med.* **375,** 2255–2262 (2016).
218. Pierce, B. G. & Weng, Z. A flexible docking approach for prediction of T cell receptor-peptide-MHC complexes. *Protein Sci.* **22,** 35–46 (2013).
219. Pierce, B. G., Vreven, T. & Weng, Z. Modeling T cell receptor recognition of CD1-lipid and MR1-metabolite complexes. *BMC Bioinformatics* **15,** 319 (2014).
220. DiMaio, F., Song, Y., Li, X., Brunner, M. J., Xu, C., Conticello, V., Egelman, E., Marlovits, T. C., Cheng, Y. & Baker, D. Atomic-accuracy models from 4.5-Å cryo-electron microscopy data with density-guided iterative local refinement. *Nat. Methods* **12,** 361–5 (2015).
221. Rubenstein, A. B., Blacklock, K., Nguyen, H., Case, D. A. & Khare, S. D. Systematic Comparison of Amber and Rosetta Energy Functions for Protein Structure Evaluation. *J. Chem. Theory Comput.* **14,** 6015–6025 (2018).
222. Leaver-Fay, A., Tyka, M., Lewis, S. M., Lange, O. F., Thompson, J. M., Jacak, R., Kaufman, K., Renfrew, P. D., Smith, C. A., Sheffler, W., Davis, I. W., Cooper, S., Treuille, A., Mandell, D. J., Richter, F., Ban, Y.-E. A., Fleishman, S. J., Corn, J. E., Kim, D. E., Berrondo, M., Mentzer, S., Popovic, Z., Havranek, J. J., Karanicolas, J., Das, R., Meiler, J., Kortemme, T., Gray, J. J., Kuhlman, B., Baker, D. & Bradley, P. ROSETTA3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. *Methods Enzymol.* **487,** 545–74 (2011).
223. Dsilva, L., Mittal, S., Koepnick, B., Flatten, J., Cooper, S. & Horowitz, S. Creating custom Foldit puzzles for teaching biochemistry. *Biochem. Mol. Biol. Educ.* **47,** 133–139 (2019).