Supplementary Material for "Antibody-Antigen Complex Modelling in the Era of Immunoglobulin Repertoire Sequencing"

Matthew I. J. Raybould,[a‡] Wing Ki Wong,[a‡] and Charlotte M. Deane[a]
[a] *Department of Statistics, University of Oxford, 24-29 St. Giles', Oxford, United Kingdom, OX1 3LB. E-mail: deane@stats.ox.ac.uk*

ANTIBODY MODELLING TOOLS

Tool	Availability	Self-reported run time	Reference
ABodyBuilder	Free	30s	1
Kotai Antibody Builder		100min	2
LYRA		35s	3
PIGS		-	4
Rosetta Antibody		60min	5
Discovery Studio (BIOVIA)	Commercial	6min	6
MOE		30-75min	7
MoFvAb		-	8
BioLuminate (Schrodinger)		-	9
SmrtMolAntibody (Macromoltek)		30min	10

PARATOPE PREDICTION TOOLS

Tool	General Description	Reference
Paratome	Sequence alignment method	11, 12
proABC	Random forest algorithm	13
Antibody i-Patch	Network approach based on binding likelihood	14
Parapred	Neural network algorithm	15

EPITOPE PREDICTION TOOLS

Tool	Description/Novelty	Epitope Category	Reference
Emini *et al.*, 1985	Surface probability profiles	Linear	16
Karplus and Schulz, 1985	B-value of C-alpha atoms	Linear	17
Parker *et al.*, 1986	Hydrophilicity scale	Linear	18
BcePred	Physicochemical properties	Linear	19
ABCPred	First neural network approach	Linear	20
PepSurf	Combinatorial phage-display libraries	Linear	21
Mapitope	Peptide libraries from mAbs	Linear	22
COBEpro	Support vector machine (SVM) for epitopic propsensity for each residues based on the fragment	Linear	23
BEST	SVM with predicted solvent accessibility and secondary structure	Linear	24
SVMTriP	SVM on tri-peptide similiarty and propensity scores	Linear	25
PepMapper	Affinity selected peptides derived from phage display, with adaptive distance threshold	Linear	26
LBtope	SVM on Chen's amino acid pair (AAP) propensities, Composition-Transition-Distribution (CTD) profile	Linear	27
APCPred	SVM on derived from amino acid anchoring pair composition (APC)	Linear	28
EPI-peptide designer	Peptide epitope designer. Find most frequent interface partners using graph analysis.	Linear	29
ElliPro	Calculate residue protrusion index, cluster neighbouring residues based on PI values	Linear / Conformational	30
Epitopia	Naive Bayes classifier	Linear / Conformational	31
EpiPred	Combines conformational matching of the antibodyantigen structures	Linear / Conformational	32
CEP	Accessibility of amino acids. First tool to predict conformational epitiopes	Conformational	33
3DEX	Physicochemical neighborhood of C-alpha/-beta atoms	Conformational	34
EPCES	Consensus scoring of propensity and physicochemical properties	Conformational	35
Rapberger *et al.*,2007	First antibody-specific epitope prediction tool; based on shape complementarity	Conformational	36
PEPITO	New physicochemical properties	Conformational	37
EpiSearch	Patch analysis that identifies contiguous clusters of residues on the surface of antigen with similar physicalchemical properties as found in phage display sequences	Conformational	38
SEPPA	Define clustering coefficient and residue neighbor of epitope patches	Conformational	39
CBTOPE	SVM on composition profile of patterns	Conformational	40
EPSVR	Support vector regression on physicochemical properties	Conformational	41
ASEP	Occurrence of residue pairs at epitope-paratope interface, followed by antibody-specific epitope propensity	Conformational	42
Bepar	Interacting residue pairs	Conformational	43
PPiPP	Neural network trained on interacting residue pairs	Conformational	44
LocaPep	Local search of epitope surface patches by residue clusters	Conformational	45
ABepar	Sequence conformational epitope prediction. coupling graph.	Conformational	46
DiscoTope 2.0	Log-odds ratio of the spatial neighbourhood and surface measures for epitope prediction	Conformational	47
BeTop	Cluster subgraphs on antigen	Conformational	48
PEASE	Residue pairing preference, with experimental input	Conformational	49
SEPIa	Random Forest (RF) / Gaussian Naive Bayes	Conformational	50
BepiPred-2.0	RF encoded with physicochemical properties	Conformational	51
Glep	Subgraph clustering for detection of epitope using SVM to detect surface patch	Conformational	52

ANTIBODY DOCKING TOOLS

Tool	Description/Novelty	Reference	Specialisation
PIPER/ClusPro Server	FFT-based, Decoys as the Reference State potential	53, 54	Ab-Ag mode
surFit	Generalized Born energy and hydration energy based on accessible surface area (GBSA)	55	Ab-Ag
SnugDock/Rosetta	Simulate induced-fit mechanism by iterating through docking optimization	56, 57	Ab-Ag
PatchDock	Connolly dot surface, shape complementarity.	58	Ab-Ag mode
FRODOCK 2.0	Spherical harmonic formulation; optimised weights for Ab-Ag	59	Ab-Ag mode
GRAMM-X	FFT-based, uses a smoothed Lennard-Jones potential on a fine grid	60	Not Ab-Ag specific
HADDOCK/PRODIGY	Semi-flexible docking with biochemical/biophysical interaction data.	61	Not Ab-Ag specific
HexServer	FFT-based optimised with GPU	62	Not Ab-Ag specific
PIER	Local statistical properties of the protein surface at the level of atomic groups	63	Not Ab-Ag specific
ProPOSE	FFT-based with side-chain flexibility	64	Not Ab-Ag specific
pyDock/FTDock	FFT-based, considers electrostatics and desolvation energy	65	Not Ab-Ag specific
SIPPER	Residue desolvation based on solvent-exposed area with the propensity-based contribution of intermolecular residue pairs.	66	Not Ab-Ag specific
SwarmDock	Flexible docking, through local docking and particle swarm optimization.	67	Not Ab-Ag specific
ZDOCK	FFT-based, shape complementarity and energy	68	Not Ab-Ag specific
InterEvDock	Incorporates evolutionary information	69	Exclude Ab-Ag

[1] J. Leem, J. Dunbar, G. Georges, J. Shi and C. M. Deane, *mAbs*, 2016, **8**, 1259–1268.
[2] K. Yamashita, K. Ikeda, K. Amada, S. Liang, Y. Tsuchiya, H. Nakamura, H. Shirai and D. M. Standley, *Bioinformatics*, 2014, **30**, 3279–3280.
[3] M. S. Klausen, M. V. Anderson, M. C. Jesperson, M. Nielsen and P. Marcatili, *Nucleic Acids Research*, 2015, **43**, W349–W355.
[4] P. Marcatili, A. Rosi and A. Tramontano, *Bioinformatics*, 2008, **24**, 1953–1954.
[5] B. D. Weitzner, J. R. Jeliazkov, S. Lyskov, N. Marze, D. Kuroda, R. Frick, J. Adolf-Bryfogle, N. Biswas, R. L. Dunbrack Jr and J. J. Gray, *Nature Protocols*, 2017, **12**, 401–416.
[6] H. Kemmish, M. Fasnacht and L. Yan, *PLoS One*, 2017, **12**, e0177923.
[7] J. K. X. Maier and P. Labute, *Proteins*, 2014, **82**, 1599–1610.
[8] A. Bujotzek, A. Fuchs, C. Qu, J. Benz, S. Klostermann, I. Antes and G. Georges, *mAbs*, 2015, **7**, 838–852.
[9] K. Zhu, T. Day, B. Warshaviak, C. Murrett, R. Friesner and D. Pearlman, *Proteins*, 2014, **82**, 1646–1655.
[10] M. Berrondo, S. Kaufmann and M. Berrondo, *Proteins*, 2014, **82**, 1636–1645.
[11] V. Kunik, B. Peters and Y. Ofran, *PLoS Computational Biology*, 2012, **8**, e1002388.
[12] V. Kunik, S. Ashkenazi and Y. Ofran, *Nucleic Acids Research*, 2012, **40**, W521–W524.
[13] P. P. Olimpieri, A. Chailyan, A. Tramontano and P. Marcatili, *Bioinformatics*, 2013, **29**, 2285–2291.
[14] K. Krawczyk, T. Baker, J. Shi and C. M. Deane, *Protein Engineering Design and Selection*, 2013, **26**, 621–629.
[15] E. Liberis, P. Veličković, P. Sormanni, M. Vendruscolo and P. Liò, *Bioinformatics*, 2018, **34**, 2944–2950.
[16] E. A. Emini, J. V. Hughes, D. S. Perlow and J. Boger, *Journal of Virology*, 1985, **55**, 836–839.
[17] P. A. Karplus and G. E. Schulz, *Naturwissenschaften*, 1985, **72**, 212–213.
[18] J. M. Parker, D. Guo and R. S. Hodges, *Biochemistry*, 1986, **25**, 5425–5432.
[19] S. Saha and G. P. S. Raghava, *Artificial Immune Systems. ICARIS 2004. Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg, 2004, pp. 197–204.
[20] S. Saha and G. P. S. Raghava, *Proteins: Structure, Function, and Bioinformatics*, 2006, **65**, 40–48.

[21] I. Mayrose, T. Shlomi, N. D. Rubinstein, J. M. Gershoni, E. Ruppin, R. Sharan and T. Pupko, *Nucleic Acids Research*, 2006, **35**, 69–78.
[22] E. M. Bublil, N. T. Freund, I. Mayrose, O. Penn, A. Roitburd-Berman, N. D. Rubinstein, T. Pupko and J. M. Gershoni, *Proteins: Structure, Function, and Bioinformatics*, 2007, **68**, 294–304.
[23] M. J. Sweredoski and P. Baldi, *Protein Engineering, Design and Selection*, 2009, **22**, 113–120.
[24] J. Gao, E. Faraggi, Y. Zhou, J. Ruan and L. Kurgan, *PLoS One*, 2012, **7**, e40104.
[25] B. Yao, L. Zhang, S. Liang and C. Zhang, *PLoS One*, 2012, **7**, e45152.
[26] W. Chen, W. W. Guo, Y. Huang and Z. Ma, *PLoS One*, 2012, **7**, e37869.
[27] H. Singh, H. R. Ansari and G. P. S. Raghava, *PLoS One*, 2013, **8**, e62216.
[28] W. Shen, Y. Cao, L. Cha, X. Zhang, X. Ying, W. Zhang, K. Ge, W. Li and L. Zhong, *BioData Mining*, 2015, **8**, 14.
[29] B. Viart, C. Dias-Lopes, E. Kozlova, C. F. Oliveira, C. Nguyen, G. Neshich, C. Chávez-Olórtegui, F. Molina and L. F. Felicori, *Bioinformatics*, 2016, **32**, 1462–1470.
[30] J. Ponomarenko, H. H. Bui, W. Li, N. Fusseder, P. E. Bourne, A. Sette and B. Peters, *BMC Bioinformatics*, 2008, **9**, 514.
[31] N. D. Rubinstein, I. Mayrose, E. Martz and T. Pupko, *BMC Bioinformatics*, 2009, **10**, 287.
[32] K. Krawczyk, X. Liu, T. Baker, J. Shi and C. M. Deane, *Bioinformatics*, 2014, **30**, 2288–2294.
[33] U. Kulkarni-Kale, S. Bhosle and A. S. Kolaskar, *Nucleic Acids Research*, 2005, **33**, W168–W171.
[34] A. Schreiber, M. Humbert, A. Benz and U. Dietrich, *Journal of Computational Chemistry*, 2005, **26**, 879–887.
[35] S. Liang, S. Liu, C. Zhang and Y. Zhou, *Proteins: Structure, Function, and Genetics*, 2007, **69**, 244–253.
[36] R. Rapberger, A. Lukas and B. Mayer, *Journal of Molecular Recognition: An Interdisciplinary Journal*, 2007, **20**, 113–121.
[37] M. J. Sweredoski and P. Baldi, *Bioinformatics*, 2008, **24**, 1459–1460.
[38] S. S. Negi and W. Braun, *Bioinformatics and Biology Insights*, 2009, **3**, 71–81.
[39] J. Sun, D. Wu, T. Xu, X. Wang, X. Xu, L. Tao, Y. X. Li and Z. W. Cao, *Nucleic Acids Research*, 2009, **37**, W612–W616.
[40] H. R. Ansari and G. P. Raghava, *Immunome Research*, 2010, **6**, 6.
[41] S. Liang, D. Zheng, D. M. Standley, B. Yao, M. Zacharias and C. Zhang, *BMC Bioinformatics*, 2010, **11**, 381.
[42] S. Soga, D. Kuroda, H. Shirai, M. Kobori and N. Hirayama, *Protein Engineering, Design and Selection*, 2010, **23**, 441–448.
[43] L. Zhao and J. Li, *BMC Structural Biology*, 2010, **10**, S6.
[44] S. Ahmad and K. Mizuguchi, *PLoS One*, 2011, **6**, e29104.
[45] L. F. Pacios, L. Tordesillas, A. Palacin, R. Sanchez-Monge, G. Salcedo and A. Diaz-Perales, *Journal of Chemical Information and Modeling*, 2011, **51**, 1465–1473.
[46] L. Zhao, L. Wong and J. Li, *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 2011, **8**, 1483–1494.
[47] J. V. Kringelum, C. Lundegaard, O. Lund and M. Nielsen, *PLoS Computational Biology*, 2012, **8**, e1002829.
[48] L. Zhao, L. Wong, L. Lu, S. C. Hoi and J. Li, *BMC Bioinformatics*, 2012, **13 (Suppl 17)**, S20.
[49] I. Sela-Culang, S. Ashkenazi, B. Peters and Y. Ofran, *Bioinformatics*, 2014, **31**, 1313–1315.
[50] G. A. Dalkas and M. Rooman, *BMC Bioinformatics*, 2017, **18**, 95.
[51] M. C. Jespersen, B. Peters, M. Nielsen and P. Marcatili, *Nucleic Acids Research*, 2017, **45**, W24–W29.
[52] L. Zhao, S. Wu, J. Jiang, W. Li, J. Luo and J. Li, *Bioinformatics*, 2018, **34**, 2061–2068.
[53] D. Kozakov, R. Brenke, S. R. Comeau and S. Vajda, *Proteins: Structure, Function, and Bioinformatics*, 2006, **65**, 392–406.
[54] D. Kozakov, D. R. Hall, B. Xia, K. A. Porter, D. Padhorny, C. Yueh, D. Beglov and S. Vajda, *Nature Protocols*, 2017, **12**, 255.
[55] N. Shimba, N. Kamiya and H. Nakamura, *Journal of Chemical Information and Modeling*, 2016, **56**, 2005–2012.
[56] A. Sircar and J. J. Gray, *PLoS Computational Biology*, 2010, **6**, e1000644.
[57] K. P. Kilambi and J. J. Gray, *Scientific Reports*, 2017, **7**, 8145.
[58] D. Schneidman-Duhovny, Y. Inbar, R. Nussinov and H. J. Wolfson, *Nucleic Acids Research*, 2005, **33**, W363–W367.
[59] E. Ramírez-Aportela, J. R. López-Blanco and P. Chacón, *Bioinformatics*, 2016, **32**, 2386–2388.
[60] A. Tovchigrechko and I. A. Vakser, *Nucleic Acids Research*, 2006, **34**, W310–W314.
[61] G. C. P. Van Zundert, J. P. G. L. M. Rodrigues, M. Trellet, C. Schmitz, P. Kastritis, E. Karaca, A. S. J. Melquiond, M. van Dijk, S. J. De Vries and A. M. J. J. Bonvin, *Journal of Molecular Biology*, 2016, **428**, 720–725.
[62] G. Macindoe, L. Mavridis, V. Venkatraman, M.-D. Devignes and D. W. Ritchie, *Nucleic Acids Research*, 2010, **38**, W445–W449.
[63] I. Kufareva, L. Budagyan, E. Raush, M. Totrov and R. Abagyan, *Proteins: Structure, Function, and Bioinformatics*, 2007, **67**, 400–417.
[64] H. Hogues, F. Gaudreault, C. R. Corbeil, C. Deprez, T. Sulea and E. O. Purisima, *Journal of Chemical Theory and Computation*, 2018, **14**, 4938–4947.
[65] T. M.-K. Cheng, T. L. Blundell and J. Fernandez-Recio, *Proteins: Structure, Function, and Bioinformatics*, 2007, **68**, 503–515.
[66] C. Pons, D. Talavera, X. de la Cruz, M. Orozco and J. Fernandez-Recio, *Journal of Chemical Information and Modeling*, 2011, **51**, 370–377.
[67] M. Torchala, I. H. Moal, R. A. Chaleil, J. Fernandez-Recio and P. A. Bates, *Bioinformatics*, 2013, **29**, 807–809.
[68] B. G. Pierce, K. Wiehe, H. Hwang, B. H. Kim, T. Vreven and Z. Weng, *Bioinformatics*, 2014, **30**, 1771–1773.
[69] C. Quignot, J. Rey, J. Yu, P. Tufféry, R. Guerois and J. Andreani, *Nucleic Acids Research*, 2018, **46**, W408–W416.