Improving the species cross-reactivity of an antibody using computational design

Christopher J. Farady [a,b,†], Benjamin D. Sellers [a,†], Matthew P. Jacobson [a,b], Charles S. Craik [a,b,*]

[a] *Graduate Group in Biophysics, University of California, San Francisco, 600 16th St., Genentech Hall, San Francisco, CA 94143-2240, USA*
[b] *Department of Pharmaceutical Chemistry, University of California, San Francisco, 600 16th St., Genentech Hall, San Francisco, CA 94143-2280, USA*

A B S T R A C T

The high degree of specificity displayed by antibodies often results in varying potencies against antigen orthologs, which can affect the efficacy of these molecules in different animal models of disease. We have used a computational design strategy to improve the species cross-reactivity of an antibody-based inhibitor of the cancer-associated serine protease MT-SP1. In silico predictions were tested in vitro, and the most effective mutation, T98R, was shown to improve antibody affinity for the mouse ortholog of the enzyme 14-fold, resulting in an inhibitor with a K_I of 340 pM. This improved affinity will be valuable when exploring the role of MT-SP1 in mouse models of cancer, and the strategy outlined here could be useful in fine-tuning antibody specificity.

© 2009 Elsevier Ltd. All rights reserved.

The ability to engineer and biochemically manipulate antibodies is critical to their utility, and has helped usher in a new wave of biological therapeutics. While bacteriophage, ribosome, and yeast display techniques have revolutionized our ability to quickly raise antibodies to a specific target, these methods are all limited by their functional library size. Considerable effort has therefore gone into improving binding beyond the initial lead antibody hit from these libraries. Iterative rounds of combinatorial techniques such as CDR walking,[1] error-prone PCR,[2] or random or site directed mutagenesis[3] have led to antibodies that regularly bind their antigens with K_D's in the picomolar range.[4] Structure-based and computational protein design techniques have also been used to streamline the antibody maturation process.[5–7]

Antibodies make outstanding therapeutic agents in part due to the high degree of specificity an antibody has for its antigen. Monoclonal antibodies often bind non-linear epitopes that depend on the precise three-dimensional arrangement of a constellation of amino acids. The specificity afforded by a monoclonal antibody has a downside, though, as it can hinder efforts to use a single antibody against species homologs. Antibodies raised against a specific human antigen often do not cross-react with non-human versions of that antigen, which can affect the choice and efficacy of pre-clinical animal models.[8,9] The same combinatorial techniques used to improve antibody affinity have been used to modify the cross-reactivity of recombinant antibodies;[9–11] here we report that computa-

tional design can be used to improve the species cross-reactivity of an antibody.

We have previously described an antibody that inhibits the extracellular serine protease membrane-type serine protease 1 (MT-SP1/matriptase). The biology of MT-SP1 is complex, but it is posited to play a role in a number of biological processes, and dysregulation of the enzyme has been shown to play a role in tumor growth and metastasis (see reviews[12,13]). The antibody, E2, was identified in a bacteriophage-displayed antibody library,[14] and is a potent inhibitor of MT-SP1, with a K_I of 12 pM.[15] Though MT-SP1 and its mouse ortholog epithin are 87% identical, E2 is a 300-fold less potent inhibitor of the mouse version of the enzyme. The basis of this difference is not apparent from biochemical and structural analyses of the MT-SP1/E2 complex. There are only three residues on the protease surface that both make contacts with the antibody and differ between the human and mouse versions of the enzyme: Ile60 is a glutamine in the mouse ortholog, Arg60c a lysine, and Tyr146 is a glutamic acid, but these residues have been shown to not be critical for inhibition.[15] In order to better understand the role MT-SP1/epithin plays in tumor progression and metastasis, mouse models of cancer[16] need to be utilized. In order to develop a version of E2 suitable for mouse experiments, computational design was used to suggest mutations predicted to improve inhibition of the mouse version of the enzyme.

Designing a species cross-reactive antibody. In silico design strategies have been used to modify protein–protein specificity,[17,18] and have successfully guided or aided in the understanding of the maturation of an antibody for an antigen.[6,7,19,20] As is frequently the case in antibody–antigen interactions, hydrogen bonding and electrostatic interactions are important contributors to the

* Corresponding author. Tel.: +1 415 476 8146.
  E-mail address: craik@cgl.ucsf.edu (C.S. Craik).
† Contributed equally to this work.

binding affinity and specificity of the interaction between MT-SP1/epithin and the antibody considered here. The MT-SP1/epithin active site prefers positively charged substrates, and the heavy and light chain CDR3 loops of E2 are positively charged.[14] To account for these electrostatic interactions, we used a molecular mechanics-based energy function, in conjunction with an implicit solvent model to treat the effects of water, to predict the effect on binding of mutations at the protease–antibody interface. This type of energy function has previously been shown to perform well in predicting mutations to increase binding affinity of antibodies,[6,7] in studying specificity of enzymes for charged metabolites,[21] and in designing enzyme active-sites.[22,23]

An estimate of the change in binding free energy upon mutation ($\Delta\Delta G_{mut}$) was calculated using the Protein Local Optimization Program (PLOP) by subtracting the calculated energies of unbound E2$_{mutated}$ and epithin species from the calculated energy of the E2$_{mutated}$/epithin complex (Fig. 1a). PLOP estimates free energy using the Optimized Potential for Liquid Simulations all atom (OPLS-AA) force field,[24–26] the Surface Generalized Born model[27] of polar solvation, an estimator for the nonpolar component of the solvation free energy,[28] and a number of correction terms as detailed in Ghosh et al.[27] and Jacobson et al.[24] Computationally estimating changes in protein conformational entropy upon binding is extremely difficult, and we do not attempt to do so. For this reason, the calculated $\Delta\Delta G_{mut}$ cannot be interpreted as a direct surrogate for the experimental change in binding affinity. The calculated values are much too large, because they do not include the entropy losses upon binding. Rather, as in previous work on specificity of proteins for small molecule substrates[21] and inhibitors,[29] the calculated values are a qualitative measure, in this case used to identify mutations that may change binding affinity and specificity.

There is currently no available structure of the epithin/E2 complex, so a homology model of epithin was created with PLOP[30] using the available MT-SP1 structure as a template. As part of the homology modeling process in PLOP, side chain rotamers are optimized[24,31] for all residues that differ between the two proteins. The epithin homology model was substituted for the protease in the MT-SP1/E2 crystal structure,[32] E2 was truncated to Fv length (ending at IleL106 and ValH111) and hydrogens were added and energy minimized.[33] A residue was selected for in silico mutation on the E2 heavy chain if at least one of its atoms was <5 Å from any atom in the three residues that differ between MT-SP1 and epithin (Gln60, Lys60c, and Glu146, Fig. 2a). The six chosen residues were ThrH28, SerH30, ThrH98, TyrH99, ProH100, and GlnH100a (Fig. 2a). The computational workflow is outlined in Figure 1b and is described as follows: (1) each of the six residue side chains were mutated to the other 18 possible amino acid side chains (excluding cysteine); (2) the side chain rotamers of residues in the interface of the complex were optimized[24,31] and then energy minimized;[33] (3) the same interface residues were energy minimized[33] in the unbound, mutated E2 and unbound epithin structures; and (4) $\Delta\Delta G_{mut}$ was calculated as described above. Most of the changes were predicted to be neutral or worsen the antibody–antigen interaction, but 8 mutations were predicted to lower the free energy of the epithin-E2 complex by varying amounts and were tested experimentally.

Testing the computational predictions. Fab constructs of the eight point mutants predicted to improve antibody binding were cloned via site-directed mutagenesis, expressed in *Escherichia coli*, and purified as previously described.[15] The IC$_{50}$'s of each point mutant were measured against epithin and MT-SP1, and relative K_I's were determined to normalize the IC$_{50}$ with respect to the protease/substrate interaction.[15] The majority of the mutations had little effect on protease inhibition (Table 1). The heavy chain P100H mutant, predicted to improve binding significantly, was deleterious to protease inhibition. This is likely because eliminating proline significantly increases the flexibility of the backbone, an effect that is

$$\Delta\Delta G_{binding}^{mutation} = \Delta G_{complex}^{mutation} - \Delta G_{E2}^{mutation} - \Delta G_{epithin}$$
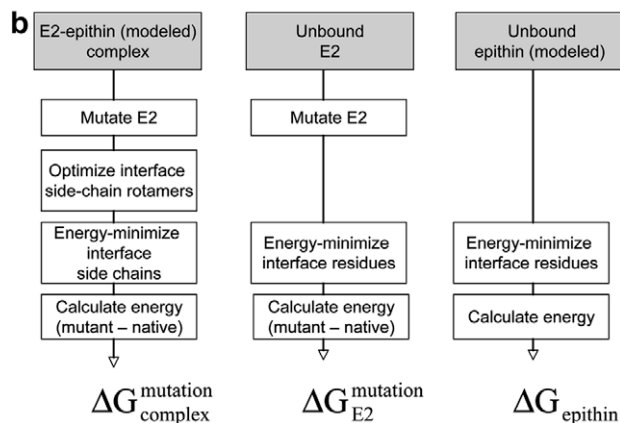

Figure 1. (a) Predicted change in binding free energy for a point mutation is estimated in silico by the calculated change in free energy upon mutation of the E2/epithin complex minus the calculated change in free energy upon mutation of the unbound E2 antibody minus the calculated free energy of epithin. (b) Details of the methods used to calculate each free energy component shown in (a) (see text).
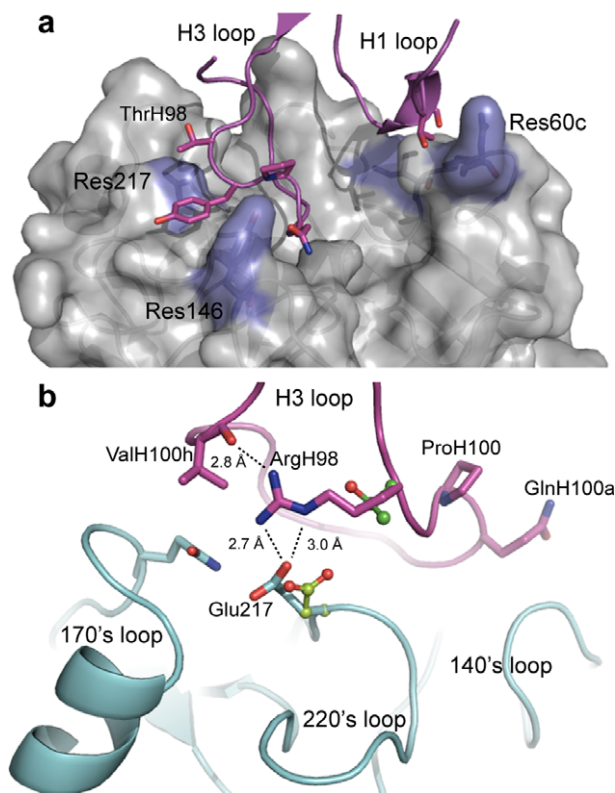

Figure 2. (a) The six residues (sticks) on the H1 and H3 CDR loops of E2 (magenta) that contact residues which are different between the human and mouse orthologs (blue) of the protease were chosen for in silico mutation. (b) Predicted effect of T98R mutation of the E2 H3 loop (magenta) on epithin (cyan) inhibition. The arginine side chain makes an intramolecular H-bond with the backbone carbonyl oxygen of ValH100h, and a hydrogen bond with the Glu217 of epithin. ThrH98 and Asp217 of MT-SP1 from the E2/MT-SP1 crystal structure are shown as ball-and-sticks. The model suggests that the T98R substitution does not affect MT-SP1 inhibition, because the Asp217 side chain cannot reach the guanidino group of ArgH98, which is locked in position by the intramolecular H-bond.

Table 1
Predicted and experimental antibody mutations

	Epithin K_i (nM)	Fold improvement[b]	$\Delta\Delta G_{exp}$[a] (kcal/mol)	$\Delta\Delta G_{predicted}$ (kcal/mol)	MT-SP1 K_i (pM)	Fold improvement[b]
E2	4.8				22.4	
T28R	3.2	1.5	−0.24	−13.2	39.3	0.6
S30G	14.0	0.3	0.63	−17.6	165.9	0.1
S30N	14.7	0.3	0.66	−11.1	155.1	0.1
T98Q	3.7	1.3	−0.16	−9.9	21.6	1.0
T98R	*0.34*	*14.1*	*−1.57*	*−21.8*	20.3	1.1
Y99S	11.5	0.4	0.52	0.12	48.6	0.5
P100H	78.5	0.1	1.66	−16.3	1346.6	0.0
Q100aV	10.3	0.5	0.45	−0.07	176.7	0.1

[a] Relative K_i's were calculated from IC_{50} values according to: $K_I = IC_{50}/(1 + [S]/\{K_M\})$. Substrate K_M's were 21 μM for epithin and 41 μM for MT-SP1. See Ref. 15 for more details.

[b] Fold improvement was determined by dividing the experimental (K_I [E2]/K_I [mutant]).

not captured in the modeling protocol. The T98R substitution, which was predicted to have the greatest effect on protease inhibition, improved the K_I of E2 for epithin from 4.8 nM to 340 pM—a 14-fold improvement. The MT-SP1/E2 crystal structure[32] provides a rationale for this improvement. The ThrH98 side chain makes no contacts with the protease, and the Oγ2 atom is solvent exposed. The model of the epithin-E2 complex suggests the T98R mutation enables the arginine side chain to extend back towards the protease, make an intramolecular H-bond with the backbone carbonyl oxygen of ValH100h, and an intermolecular H-bond with the carboxylate group of Glu217 on the protease (Fig. 2b). The 14-fold improvement in K_I shown by this construct corresponds to a free energy gain of 1.6 kcal/mol, which is roughly the strength of one strong or two moderately strong H-bonds. This model also suggests why the T98R mutation has little effect on MT-SP1 inhibition. The human ortholog has an Asp residue at position 217, which, because its side chain is one atom shorter, cannot extend far enough to make this interaction. As a result, the threonine to arginine substitution allows a side chain that plays a minimal role in protease inhibition to be modified to broaden the species cross-reactivity of the E2 antibody.

Our broadening of the species specificity of E2 highlights a number of reasons the antibody–antigen interaction is ideally suited for modification by computational design. When bound to proteins, antibodies often have more than 800 square angstroms of buried surface area, and are dominated by binding 'hot-spots',[34] which allows for the possibility of optimization at a number of discrete locations without destroying binding. Mutational studies on the MT-SP1/E2 complex suggested the antibody–antigen complex is dependent on interactions at ArgH100b and portions of the H2 and H3 loops that 'grabbed' the 90s loop on the protease,[15] so we did not computationally screen these areas. Furthermore, the energetics of antibody–antigen complexes tend to be dependent on side chain-side chain interactions more than main chain interactions.[35] Side-chain interactions can be easier to model, particularly when the protein backbone is relatively rigid. This is illustrated by our success modeling the effects of mutating side chain interactions at residue Thr98, and the difficulty experienced modeling Pro99, the only residue that was modified which significantly reduced inhibition. Also, the shape complementarity (as defined by the geometric match at a protein–protein interface) of antibody–antigen interactions tends to be worse than other well characterized protein–protein interactions,[36] and antibody CDR sequences are biased towards specific residues such as tyrosine, tryptophan, and aspartic acid.[37] This is likely a function of the inherent flexibility antibodies must have to bind to new and different antigens, and is clearly an efficient way to rapidly develop binders to many different targets, but likely leaves a portion of sequence and conformational space underexplored, and suggests there are a number

of opportunities for computational modeling to predict mutations that could fine-tune the interaction.

In conclusion, computational design was used to predict a suite of mutations that could improve the species cross-specificity of an inhibitory antibody of the cancer-associated serine protease MT-SP1. The substitution of an arginine for ThrH98 improved the affinity of the antibody for epithin by an order of magnitude, and had no effect on inhibition of the human ortholog. In conjunction with other similar published studies,[6,7] the results presented here suggest that computational strategies can help guide the fine-tuning of antibody–antigen interactions, and streamline the process by which therapeutic antibodies are investigated in different animal models.

Acknowledgements

References and notes

1. Barbas, C. F., 3rd; Hu, D.; Dunlop, N.; Sawyer, L.; Cababa, D.; Hendry, R. M.; Nara, P. L.; Burton, D. R. Proc. Natl. Acad. Sci. U.S.A. **1994**, 91, 3809.
2. Razai, A.; Garcia-Rodriguez, C.; Lou, J.; Geren, I. N.; Forsyth, C. M.; Robles, Y.; Tsai, R.; Smith, T. J.; Smith, L. A.; Siegel, R. W.; Feldhaus, M.; Marks, J. D. J. Mol. Biol. **2005**, 351, 158.
3. Ho, M.; Kreitman, R. J.; Onda, M.; Pastan, I. J. Biol. Chem. **2005**, 280, 607.
4. Luginbuhl, B.; Kanyo, Z.; Jones, R. M.; Fletterick, R. J.; Prusiner, S. B.; Cohen, F. E.; Williamson, R. A.; Burton, D. R.; Pluckthun, A. J. Mol. Biol. **2006**, 363, 75.
5. Barderas, R.; Desmet, J.; Timmerman, P.; Meloen, R.; Casal, J. I. Proc. Natl. Acad. Sci. U.S.A. **2008**, 105, 9029.
6. Clark, L. A.; Boriack-Sjodin, P. A.; Eldredge, J.; Fitch, C.; Friedman, B.; Hanf, K. J.; Jarpe, M.; Liparoto, S. F.; Li, Y.; Lugovskoy, A.; Miller, S.; Rushe, M.; Sherman, W.; Simon, K.; Van Vlijmen, H. Protein Sci. **2006**, 15, 949.
7. Lippow, S. M.; Wittrup, K. D.; Tidor, B. Nat. Biotechnol. **2007**, 25, 1171.
8. Pegram, M.; Ngo, D. Adv. Drug Deliv. Rev. **2006**, 58, 723.
9. Werther, W. A.; Gonzalez, T. N.; O'Connor, S. J.; McCabe, S.; Chan, B.; Hotaling, T.; Champe, M.; Fox, J. A.; Jardieu, P. M.; Berman, P. W.; Presta, L. G. J. Immunol. **1996**, 157, 4986.
10. Garcia-Rodriguez, C.; Levy, R.; Arndt, J. W.; Forsyth, C. M.; Razai, A.; Lou, J.; Geren, I.; Stevens, R. C.; Marks, J. D. Nat. Biotechnol. **2007**, 25, 107.
11. Liang, W. C.; Wu, X.; Peale, F. V.; Lee, C. V.; Meng, Y. G.; Gutierrez, J.; Fu, L.; Malik, A. K.; Gerber, H. P.; Ferrara, N.; Fuh, G. J. Biol. Chem. **2006**, 281, 951.
12. Bugge, T. H.; List, K.; Szabo, R. Front Biosci. **2007**, 12, 5060.
13. Darragh, M. R.; Bhatt, A. S.; Craik, C. S. Front Biosci. **2008**, 13, 528.
14. Sun, J.; Pons, J.; Craik, C. S. Biochemistry **2003**, 42, 892.
15. Farady, C. J.; Sun, J.; Darragh, M. R.; Miller, S. M.; Craik, C. S. J. Mol. Biol. **2007**, 369, 1041.
16. Welm, A. L.; Sneddon, J. B.; Taylor, C.; Nuyten, D. S.; van de Vijver, M. J.; Hasegawa, B. H.; Bishop, J. M. Proc. Natl. Acad. Sci. U.S.A. **2007**, 104, 7570.
17. Joachimiak, L. A.; Kortemme, T.; Stoddard, B. L.; Baker, D. J. Mol. Biol. **2006**, 361, 195.

C. J. Farady et al./Bioorg. Med. Chem. Lett. 19 (2009) 3744–3747

3747

18. Shifman, J. M.; Mayo, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13274.
19. Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14330.
20. Sivasubramanian, A.; Maynard, J. A.; Gray, J. J. *Proteins* **2008**, *70*, 218.
21. Kalyanaraman, C.; Bernacki, K.; Jacobson, M. P. *Biochemistry* **2005**, *44*, 2059.
22. Chakrabarti, R.; Klibanov, A. M.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12035.
23. Chakrabarti, R.; Klibanov, A. M.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10153.
24. Jacobson, M. P.; Kaminski, G. A.; Friesner, R. A.; Rapp, C. S. *J. Phys. Chem. B* **2002**, *106*, 11673.
25. Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *Q. Rev. Biophys.* **1993**, *26*, 49.
26. Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474.
27. Ghosh, A.; Rapp, C. S.; Friesner, R. A. *J. Phys. Chem. B* **1998**, *102*, 10983.
28. Gallicchio, E.; Zhang, L. Y.; Levy, R. M. *J. Comp. Chem.* **2002**, *23*, 517.
29. Huang, N.; Kalyanaraman, C.; Bernacki, K.; Jacobson, M. P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5166.
30. Kenyon, V.; Chorny, I.; Carvajal, W. J.; Holman, T. R.; Jacobson, M. P. *J. Med. Chem.* **2006**, *49*, 1356.
31. Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. *J. Mol. Biol.* **2002**, *320*, 597.
32. Farady, C. J.; Egea, P. F.; Schneider, E. L.; Darragh, M. R.; Craik, C. S. *J. Mol. Biol.* **2008**, *380*, 351.
33. Zhu, K.; Shirts, M. R.; Friesner, R. A.; Jacobson, M. P. *J. Chem. Theory Comput.* **2007**, *3*, 640.
34. Dall'Acqua, W.; Goldman, E. R.; Eisenstein, E.; Mariuzza, R. A. *Biochemistry* **1996**, *35*, 9667.
35. Jackson, R. M. *Protein Sci.* **1999**, *8*, 603.
36. Lawrence, M. C.; Colman, P. M. *J. Mol. Biol.* **1993**, *234*, 946.
37. Fellouse, F. A.; Wiesmann, C.; Sidhu, S. S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12467.
38. DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific: Palo Alto, CA. USA, 2002.