Mutation Effects on Remote Epitopes of An Autoantigen Decoded with Simulated B-Factors

Yuan-Ping Pang[1*], Marta Casal Moura[2], Gwen E. Thompson[2], Amber M. Hummel[2], Darlene R. Nelson[2], Dieter E. Jenne[3], and Ulrich Specks[2*]

Proteinase 3 (PR3) is a neutrophil serine protease targeted by anti-neutrophil cytoplasmic antibodies (ANCAs) in the autoimmune disease granulomatosis with polyangiitis (GPA)[1-5]. PR3 mutants were developed to investigate how PR3 interacts with the ANCAs and whether the interactions can be intervened by therapeutics. One mutant with a Ser195Ala mutation (iPR3-Val^{103}) recognized as many ANCAs as wild-type PR3 (PR3-Val^{103})[6-9], indicating that PR3-Val^{103} and iPR3-Val^{103} have equivalent ANCA-binding capabilities. A triple mutant of the latter (iHm5-Val^{103}) bound a monoclonal antibody (moANCA518) from a patient with GPA on an epitope remote from the mutation sites. Unexpectedly, the corresponding epitope of iPR3-Val^{103} was inaccessible to moANCA518 under the same experimental conditions[10,11]. These observations demonstrate that a latent epitope of PR3 can be activated surprisingly by remote mutations in PR3[11]. Here we report a comparative analysis of simulated B-factors (i.e., measurements of local mobility) of PR3-Val^{103}, iPR3-Val^{103}, and iHm5-Val^{103}, demonstrating that the binding of moANCA518 to iHm5-Val^{103} is enabled by an increase in main-chain flexibility in the latent epitope caused by the remote muta-

[1]Computer-Aided Molecular Design Laboratory, Mayo Clinic, Rochester, MN 55905, USA; [2]Thoracic Disease Research Unit, Mayo Clinic, Rochester, MN 55905, USA; [3]Institute of Lung Biology and Disease, Helmholtz-Center, Munich, and Max Planck Institute of Neurobiology, Planegg-Martinsried, Germany. *E-mail: pang@mayo.edu; specks.ulrich@mayo.edu

tions in iHm5-Val^{103}. This epitope activation—achieved in vitro by remote mutations as we demonstrated or in vivo conceivably by remote protein•protein interactions[12] or remote polymorphisms—may be a fundamental feature of antibody-mediated autoimmune diseases. Rigidifying B-cell epitopes on an autoantigen with therapeutics designed using the B-factor analysis disclosed here may lead to effective treatments for these autoimmune diseases by making the autoantigen inaccessible to existing autoantibodies. This analysis may also be used to predict and characterize epitopes and remote mutation effects.

To investigate how the structure of PR3 affects its capability to bind ANCAs, we developed a human PR3 mutant (iPR3-Val^{103}) with a Ser195Ala mutation that perturbs the catalytic triad comprising Asp102, His57, and Ser195 and thereby disables catalytic functioning in PR3[6-9]. This mutation is close to Epitope 5 of PR3 and remote from Epitopes 1, 3, and 4 (Fig. 1)[10]. Our finding that iPR3-Val^{103} recognized as many ANCA serum samples from patients as wild-type PR3 (PR3-Val^{103}) in both immunofluorescence assay and ELISA indicates the equivalence between iPR3-Val^{103} and PR3-Val^{103} in terms of their ANCA-binding capabilities[7].

We also developed iHm5-Val^{103}—a chimeric triple mutant of iPR3-Val^{103} in which Ala146, Trp218, and Leu223 are replaced by Thr146, Arg218, and Gln223 from mouse PR3—to demonstrate altered binding of ANCAs to Epitope 5 (and possibly Epitope 1 but not Epitopes 3 and 4) of the mutant because Trp218 and Leu223 reside in Epitope 5 and Ala146 is in Epitope 1 (Fig. 1)[10]. Unexpectedly, a monoclonal ANCA (moANCA518) cloned from a patient with GPA was found to bind to Epitope 3 of iHm5-Val^{103}, which is distal to the three mutation sites[11]. By contrast, moANCA518 did not bind to Epitope 3 of iPR3-Val^{103} under the same experimental conditions[11]. More importantly, the majority of serum samples from patients with GPA showed similar preferential binding to Epitope 3 of iHm5-Val^{103} compared with iPR3-Val^{103} [11]. These observations indicate that Epitope 3, a mutation-free epitope of PR3, is latent in iPR3-Val^{103} but active in iHm5-Val^{103} for ANCA binding. It also indicates that a latent epitope of PR3 is activated by remote muta-

tions, which is akin to our reported finding that a monoclonal antibody (MCPR3-7) allosterically interfered with the activity of PR3[12].

These clinical observations raised a mechanistic question: How can a latent epitope, such as the one in iPR3-Val^{103}, be activated by topologically distal mutations, such as those in iHm5-Val^{103}? Answering this question may help explain how MCPR3-7 allosterically inhibits PR3 activity because MCPR3-7 bind to the region in PR3 that encompasses the three chimeric mutation sites[12]. More importantly, understanding how PR3 activity can be affected by remote mutations or remote protein•protein interactions may offer insights into the causes of GPA and other antibody-mediated autoimmune diseases and the development of new treatments for these diseases.

Accordingly, we developed computer models of PR3-Val^{103}, iPR3-Val^{103}, and iHm5-Val^{103} to try to understand how mutations of these variants affect the ANCA-binding capabilities of the four reported epitopes of PR3[10]. These models were derived from molecular dynamics (MD) simulations using our published forcefield and simulation protocol[13]. The initial conformations of the three variants used in these simulations were derived via mutation(s) from the PR3-Ile^{103} crystal structure[14] because experimentally determined structures of these variants have been unavailable to date. Although local differences in main-chain conformations of two surface loops between iHm5-Val^{103} and PR3-Val^{103} (or iHm5-Val^{103} and iPR3-Val^{103}) were observed (Fig. S1), the overall conformations of the three variants resembled one another according to the alpha carbon (Cα) root mean square deviations of ≤1.63 Å (Table S1). Given these conformational properties, we could not determine how mutations of these variants affect the ANCA-binding capabilities of the PR3 epitopes, primarily because these surface loops are highly flexible and lack the time dimension that is required for immunological function analysis[15].

To take the time dimension into account, we turned our attention to the dynamic properties of the PR3 variants. It is well-known that a folded protein is fluid-like with fluctuations in atomic position on the picosecond timescale and that the dynamics of these atomic displacements are

dominated by collisions with neighboring atoms involving reorientation of side chains or localized portions of the backbone[16]. Two seminal studies have also shown that the crystallographically determined high B-factors of a protein fragment are linked to the antigenicity of the fragment[17,18]. This link indicates that the crystallographically determined B-factor—defined as $8\pi^2\langle u^2\rangle$ to reflect the displacement u of the atom from its mean position, thermal motions, local mobility, or the uncertainty of the atomic mean position[19-27]—can be used to aid the identification and characterization of antigen epitopes.

However, the crystallographically determined B-factor of an atom reflects not only the thermal motion or local mobility of the atom but also conformational and static lattice disorders of the atom, and even the refinement error in determining the mean position of the atom[22,24,26,28]. Therefore, using crystallographically determined B-factors to investigate epitopes requires the comparison of B-factors of different crystal structures of the same protein, which are in different space groups and obtained with different refinement procedures at different resolutions, in order to identify the B-factors that reflect the local mobility of the protein[28].

This requirement can be avoided by using simulated B-factors derived from MD simulations on a picosecond timescale because simulated B-factors are devoid of refinement errors and conformational and static lattice disorders. In addition, local motions, such as those of backbone N–H bonds, occur on the order of tens or hundreds of picoseconds[29].

In this context, we calculated the Cα B-factors of PR3-Val^{103}, iPR3-Val^{103}, and iHm5-Val^{103} from MD simulations on a 50-ps timescale using our published forcefield[13] and method[30]. The mean Cα B-factors of PR3-Val^{103}, iPR3-Val^{103}, and iHm5-Val^{103} were 6.84 Å^2 (95%CI: 6.75–6.94 Å^2), 6.91 Å^2 (95%CI: 6.82–7.00 Å^2), and 7.13 Å^2 (95%CI: 7.03–7.24 Å^2), respectively, wherein 95%CI is the abbreviation of 95% confidence interval. Given these findings, we concluded that any surface loop could be considered highly mobile and hence potentially antigenic if the mean Cα B-factor of the loop were >9.00 Å^2. According to this criterion, PR3-Val^{103} has 10 poten-

tially antigenic surface loops, and iPR3-Val[103] and iHm5-Val[103] have 11 each (Fig. 2). Consistent with the two seminal reports[17,18], all of these potentially antigenic loops identified a priori by using simulated B-factors fall within all four known epitopes of PR3[10], demonstrating a clear association between a loop with a high mean simulated Cα B-factor and the experimentally determined antigenicity of the loop.

Further, we found that the Ser195Ala mutation caused no significant reduction in the mean Cα B-factor of any of the 10 potentially antigenic surface loops in PR3-Val[103] (Fig. 2A). This finding implies that the Ser195Ala mutation would not impair the ANCA-binding capability of any of the four epitopes of iPR3-Val[103], and it explains our reported observation that iPR3-Val[103] recognizes as many ANCA serum samples as PR3-Val[103] does[7].

We also found the mean Cα B-factors of Loop 3B in iPR3-Val[103] (possessing Ala146, Trp218, and Leu223) and iHm5-Val[103] (possessing Thr146, Arg218, and Gln223) to be 6.9 $Å^2$ (95%CI: 6.8–7.0 $Å^2$) and 12.8 $Å^2$ (95%CI: 12.3–13.2 $Å^2$), respectively (Fig. 2B). According to the antigenicity criterion of 9.00 $Å^2$, these means suggest that the three chimeric mutations make Loop 3B (a mutation-free loop) more mobile in iHm5-Val[103], despite large separations between Epitope 3 of PR3 and the chimeric mutation sites (~32 Å, ~32 Å, and ~31 Å from the Cα atom of Gln122 in Epitope 3 to the Cα atoms of Ala146, Trp218, and Leu223, respectively, at the chimeric mutation sites). Therefore, Epitope 3 of iHm5-Val[103] would bind ANCAs, whereas the ANCA-binding capability of Epitope 3 of iPR3-Val[103] would be rather limited. These findings also explain why moANCA518 and other ANCAs from patients with GPA bind to Epitope 3 of iHm5-Val[103] rather than iPR3-Val[103] in immunoassays[11].

The above analyses of simulated B-factors and our clinical observations[7,11] reveal a previously unrecognized mechanism for epitope activation: Remote mutations can increase the local mobility (i.e., main-chain flexibility) of a latent epitope of an antigen, which facilitates the conformational adaptation required for antibody binding and thereby activate the latent epitope. In the

same vein, it is plausible that the binding of MCPR3-7 to PR3 modulates the main-chain flexibility of a topographically remote site of PR3, which contributes to the reported allosteric inhibition of PR3 by MCPR3-7[12]. This type of exquisite epitope activation or inhibition—achieved either in vitro by remote mutations as we demonstrated or in vivo conceivably by remote protein•protein interactions or remote polymorphisms—may be a fundamental feature of antibody-mediated autoimmune diseases. This feature has implications for the development of novel treatments that aim to disrupt the pathogenic autoantibody•autoantigen interactions in these autoimmune diseases by reducing the main-chain flexibility of a target B-cell epitope using chemical or biological therapeutics, which will be designed with the simulated B-factor analysis protocol disclosed here. This protocol may also be used to predict and characterize B-cell epitopes and effects of remote mutations.

1. Wegener, F. Über generalisierte septische Gefäßerkrankungen. *Verh. Dtsch. Pathol. Ges.* **29**, 202–209 (1936).
2. Fahey, J. L., Leonard, E., Churg, J. & Godman, G. Wegener's granulomatosis. *Am. J. Med.* **17**, 168–179 (1954).
3. Fauci, A. S. & Wolff, S. M. Wegener's granulomatosis: studies in eighteen patients and a review of the literature. *Medicine (Baltimore)* **52**, 535–561 (1973).
4. Jenne, D. E., Tschopp, J., Ludemann, J., Utecht, B. & Gross, W. L. Wegener's autoantigen decoded. *Nature* **346**, 520 (1990).
5. Gross, W. L. & Csernok, E. Immunodiagnostic and pathophysiologic aspects of antineutrophil cytoplasmic antibodies in vasculitis. *Curr. Opin. Rheumatol.* **7**, 11–19 (1995).
6. Specks, U., Fass, D. N., Fautsch, M. P., Hummel, A. M. & Viss, M. A. Recombinant human proteinase 3, the Wegener's autoantigen, expressed in HMC-1 cells is enzymatically active and recognized by c-ANCA. *FEBS Lett.* **390**, 265–270 (1996).

7. Sun, J. *et al.* Capture-ELISA based on recombinant PR3 is sensitive for PR3-ANCA testing and allows detection of PR3 and PR3-ANCA/PR3 immunecomplexes. *J. Immunol. Methods* **211**, 111–123 (1998).
8. Sun, J. *et al.* A proportion of proteinase 3 (PR3)-specific anti-neutrophil cytoplasmic antibodies (ANCA) only react with PR3 after cleavage of its N-terminal activation dipeptide. *Clin. Exp. Immunol.* **114**, 320–326 (1998).
9. Specks, U. What you should know about PR3-ANCA. Conformational requirements of proteinase 3 (PR3) for enzymatic activity and recognition by PR3-ANCA. *Arthritis Res.* **2**, 263–267 (2000).
10. Silva, F., Hummel, A. M., Jenne, D. E. & Specks, U. Discrimination and variable impact of ANCA binding to different surface epitopes on proteinase 3, the Wegener's autoantigen. *J. Autoimmun.* **35**, 299–308 (2010).
11. Casal Moura, M. *et al.* Preferential binding of anti-neutrophil cytoplasmic antibodies to an unexpected epitope of a chimeric proteinase 3 mutant. *BioRxiv* (2019) doi: https://doi.org/10.1101/549063.
12. Hinkofer, L. C. *et al.* A monoclonal antibody (MCPR3-7) interfering with the activity of proteinase 3 by an allosteric mechanism. *J. Biol. Chem.* **288**, 26635–26648 (2013).
13. Pang, Y.-P. FF12MC: a revised AMBER forcefield and new protein simulation protocol. *Proteins* **84**, 1490–1516 (2016).
14. Fujinaga, M., Chernaia, M. M., Halenbeck, R., Koths, K. & James, M. N. The crystal structure of PR3, a neutrophil serine proteinase antigen of Wegener's granulomatosis antibodies. *J. Mol. Biol.* **261**, 267–278 (1996).
15. Van Regenmortel, M. H. Reductionism and the search for structure-function relationships in antibody molecules. *J. Mol. Recognit.* **15**, 240–247 (2002).

16. McCammon, J. A., Gelin, B. R. & Karplus, M. Dynamics of folded proteins. *Nature* **267**, 585–590 (1977).

17. Westhof, E. *et al.* Correlation between segmental mobility and the location of antigenic determinants in proteins. *Nature* **311**, 123–126 (1984).

18. Tainer, J. A. *et al.* The reactivity of anti-peptide antibodies is a function of the atomic mobility of sites in a protein. *Nature* **312**, 127–134 (1984).

19. Debye, P. Interference of x rays and heat movement. *Ann. Phys.* **43**, 49–95 (1913).

20. Waller, I. On the effect of thermal motion on the interference of X-rays. *Z. Phys.* **17**, 398–408 (1923).

21. Willis, B. T. M. & Pryor, A. W. *Thermal vibrations in crystallography*. (Cambridge University Press, 1975).

22. Yu, H. A., Karplus, M. & Hendrickson, W. A. Restraints in temperature-factor refinement for macromolecules: an evaluation by molecular dynamics. *Acta Crystallogr., Sect. B: Struct. Sci.* **41**, 191–201 (1985).

23. Kidera, A. & Go, N. Normal mode refinement: crystallographic refinement of protein dynamic structure. 1. Theory and test by simulated diffraction data. *J. Mol. Biol.* **225**, 457–475 (1992).

24. McRee, D. E. *Practical protein crystallography*. 98 (Academy Press, 1993).

25. Trueblood, K. N. *et al.* Atomic displacement parameter nomenclature: report of a subcommittee on atomic displacement parameter nomenclature. *Acta Crystallogr., Sect. A* **52**, 770–781 (1996).

26. Tronrud, D. E. Knowledge-based B-factor restraints for the refinement of proteins. *J. Appl. Crystallogr.* **29**, 100-104 (1996).

27. Garcia, A. E., Krumhansl, J. A. & Frauenfelder, H. Variations on a theme by Debye and Waller: from simple crystals to proteins. *Proteins* **29**, 153–160 (1997).

28. Artymiuk, P. J. *et al.* Crystallographic studies of the dynamic properties of lysozyme. *Nature* **280**, 563–568 (1979).
29. Morin, S. A practical guide to protein dynamics from ^{15}N spin relaxation in solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **59**, 245–262 (2011).
30. Pang, Y.-P. Use of multiple picosecond high-mass molecular dynamics simulations to predict crystallographic B-factors of folded globular proteins. *Heliyon* **2**, e00161 (2016).

Acknowledgements This work was supported by the US Army Research Office (W911NF-16-1-0264; to Y.-P.P.), the Connor Group Foundation (to U.S.), the Mayo Foundation for Medical Education and Research (to Y.-P.P. and U.S.), and the European Union's Horizon 2020 research and innovation program under grant agreement No 668036 (RELENT; to D.E.J.). Responsibility for the information and views in this study lies entirely with the authors.

Author contributions D.R.N. and U.S. initiated the collaborative project. U.S. and D.E.J. designed the PR3 variants and the ANCA-binding experiments. M.C.M., G.E.T., A.M.H., and D.R.N. performed the ANCA-binding experiments. Y.-P.P. conceived the idea of using simulated B-factors to computationally study B-cell epitopes and performed the B-factor calculations. Y.-P.P., U.S., and D.E.J. wrote the manuscript. Y.-P.P., M.C.M., G.E.T., D.E.J., and U.S. contributed to revisions of the manuscript.

Competing interests The authors have no conflicts of interest.

Additional information Tables S1–S2 and Figure S1

METHODS

Initial conformation preparation. The initial conformation of PR3-Ile[103] (residues 16–239; truncated for atomic charge neutrality) was taken from the crystal structure of PR3[14]. The initial conformations of the corresponding PR3-Val[103] and iPR3-Val[103] (residues 16–239) were taken from the initial PR3-Ile[103] conformation with mutations of Ile103Val alone and Ile103Val together with Ser195Ala, respectively. The initial conformation of iHm5-Val[103] (residues 16–238; truncated for atomic charge neutrality) was taken from the initial PR3-Ile[103] conformation with mutations of Ala146Thr, Trp218Arg, Leu223Gln, Ile103Val, and Ser195Ala. The crystallographically determined water molecules with residue identifiers of 246–249, 257–259, 261–263, 268, 270, 274–276, 279, 280, 291, 292, 296, 298, 307, 309, and 317 were included in all four conformations. The AMBER residue names of ASP, GLU, ARG, LYS, HID, and CYX were used for all Asp, Glu, Arg, Lys, His, and Cys residues, respectively. All initial conformations were refined via energy minimization using the SANDER module of AMBER 11 (University of California, San Francisco) and forcefield FF12MClm[13] with a dielectric constant of 1.0, a cutoff of 30.0 Å for nonbonded interactions, and 200 cycles of steepest descent minimization followed by 100 cycles of conjugate gradient minimization.

MD simulations. Each of the four energy-minimized conformations described above was solvated with 5578 (for iHm5-Val[103]) or 5536 (for all other variants) TIP3P[31] water molecules (using "solvatebox PR3 TIP3BOX 8.2") and then energy-minimized for 100 cycles of steepest descent minimization followed by 900 cycles of conjugate gradient minimization using SANDER of AMBER 11 to remove close van der Waals contacts. The initial solvation box size was 58.268 × 68.409 × 65.657 $Å^3$ (for iHm5-Val[103]) or 67.337 × 66.050 × 58.335 $Å^3$ (for all other variants). The resulting system was heated from 5 K to 340 K at a rate of 10 K/ps under constant temperature and constant volume, then equilibrated for 10^6 timesteps under a constant temperature of 340 K and a constant pressure of 1 atm using the isotropic molecule-based scaling. Finally, 20 distinct,

independent, unrestricted, unbiased, isobaric–isothermal, 316-ns MD simulations of the equilibrated system with forcefield FF12MClm[13] were performed using PMEMD of AMBER 11 with a periodic boundary condition at 340 K and 1 atm. The 20 unique seed numbers for initial velocities of the 20 simulations were taken from Ref.[32]. All simulations used (*i*) a dielectric constant of 1.0, (*ii*) the Berendsen coupling algorithm[33], (*iii*) the particle mesh Ewald method to calculate electrostatic interactions of two atoms at a separation of >8 Å[34], (*iv*) Δt = 1.00 fs of the standard-mass time[13], (*v*) the SHAKE-bond-length constraint applied to all bonds involving hydrogen, (*vi*) a protocol to save the image closest to the middle of the "primary box" to the restart and trajectory files, (*vii*) a formatted restart file, (*viii*) the revised alkali and halide ion parameters[35], (*ix*) a cutoff of 8.0 Å for nonbonded interactions, (*x*) a uniform 10-fold reduction in the atomic masses of the entire simulation system (both solute and solvent), and (*xi*) default values of all other inputs of the PMEMD module. The forcefield parameters of FF12MClm are available in the Supporting Information of Ref.[36]. All simulations were performed on a cluster of 100 12-core Apple Mac Pros with Intel Westmere (2.40/2.93 GHz).

Cα B-factor calculation. In a two-step procedure using PTRAJ of AmberTools 1.5, the B-factors of Cα atoms in PR3 were calculated from all conformations saved at every 10^3 timesteps during 20 simulations of the protein using the simulation conditions described above except that (*i*) the atomic masses of the entire simulation system (both solute and solvent) were uniformly increased by 100-fold relative to the standard atomic masses, (*ii*) the simulation temperature was lowered to 300 K, and (*iii*) the simulation time was reduced to 500 ps. The first step was to align all saved conformations onto the first saved conformation to obtain an average conformation using the root mean square fit of all Cα atoms. The second step was to perform root mean square fitting of all Cα atoms in all saved conformations onto the corresponding atoms of the average conformation. The Cα B-factors were then calculated using the "atomicfluct" command in

PTRAJ. For each protein, the calculated B-factor of any atom in Table S2 was the mean of all B-factors of the atom derived from 20 simulations of the protein. The standard error (SE) of a B-factor was calculated according to Eq. 2 of Ref.[37]. The SE of the average Cα B-factor of each PR3 variant was calculated according to the standard method for propagation of errors of precision[38]. The 95% confidence interval (95%CI) of the average Cα B-factor was obtained according to the formula mean ± 1.96 × SE because the sample size of each PR3 variant exceeded 100.

Conformational cluster analysis and root mean square deviation calculation. The conformational cluster analyses were performed using CPPTRAJ of AmberTools 16 with the average-linkage algorithm[39], epsilon of 3.0 Å, and root mean square coordinate deviation on all Cα atoms of the proteins (see Table S1). Cα root mean square deviations (CαRMSDs) were manually calculated using ProFit V2.6 (http://www.bioinf.org.uk/software/profit/). The first unit of the crystal structure of the PR3 tetramer and the average conformation (without energy minimization) of the most populated cluster were used for the CαRMSD calculation.

31. Jorgensen, W. L., Chandreskhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
32. Pang, Y.-P. Use of 1–4 interaction scaling factors to control the conformational equilibrium between α-helix and β-strand. *Biochem. Biophys. Res. Commun.* **457**, 183–186 (2015).
33. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Di Nola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
34. Darden, T. A., York, D. M. & Pedersen, L. G. Particle mesh Ewald: an N log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

35. Joung, I. S. & Cheatham, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).

36. Pang, Y.-P. Low-mass molecular dynamics simulation for configurational sampling enhancement: more evidence and theoretical explanation. *Biochem. Biophys. Rep.* **4**, 126–133 (2015).

37. Pang, Y.-P. At least 10% shorter C–H bonds in cryogenic protein crystal structures than in current AMBER forcefields. *Biochem. Biophys. Res. Commun.* **458**, 352–355 (2015).

38. Andraos, J. On the propagation of statistical errors for a function of several variables. *J. Chem. Educ.* **73**, 150–154 (1996).

39. Shao, J., Tanner, S. W., Thompson, N. & Cheatham III, T. E. Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.* **3**, 2312–2334 (2007).

Figure legends

Fig. 1. Front and back views of PR3 depicting its four known epitopes, each comprising multiple surface loops with high Cα B-factors derived from simulations. L1A: Loop 1A of residues 36–38C; L1B: Loop 1B of residues 145–151; L1C: Loop 1C of residues 75–79; L3A: Loop 3A of residues 110–117; L3B: Loop 3B of residues 124–133; L3C: Loop 3C of residues 202–204; L4A: Loop 4A of residues 59–63C; L4B: Loop 4B of residues 92–99; L5A: Loop 5A of residues 165–178; L5B: Loop 5B of residues 186–187; L5C: Loop 5C of residues 192–194; and L5D: Loop 5D of residues 219–224; wherein the residue numbering here is identical to that of the PR3 crystal structure (PDB ID: 1FUJ).

Fig. 2. Simulated Cα B-factors of PR3-Val^{103}, iPR3-Val^{103}, and iHm5-Val^{103}. The simulated mean Cα B-factors of PR3-Val^{103}, iPR3-Val^{103}, and iHm5-Val^{103} are 6.84 Å^2 (95%CI: 6.75–6.94 Å^2; labeled as avg-PR3-Val^{103}), 6.91 Å^2 (95%CI: 6.82–7.00 Å^2; labeled as avg-iPR3-Va^{103}), and 7.13 Å^2 (95%CI: 7.03–7.24 Å^2; labeled as avg-iHm5-Val^{103}), respectively, wherein 95%CI is the abbreviation of 95% confidence interval. The simulated Cα B-factors were plotted using the human PR3 sequence (NCBI P24158.3) numbering because the PR3 crystal structure numbering is discontinuous. Therefore, the following loop residues are defined using the PR3 crystal structure numbering followed by the NCBI P24158.3 numbering in parenthesis. L1A: Loop 1A of residues 36–38C(48–52); L1B: Loop 1B of residues 145–151(161–166); L1C: Loop 1C of residues 75–79(92–96); L3A: Loop 3A of residues 110–117(126–133); L3B: Loop 3B of residues 124–133(140–149); L3C: Loop 3C of residues 202–204(210–212); L4A: Loop 4A of residues 59–63C(73–80); L4B: Loop 4B of residues 92–99(108–115); L5A: Loop 5A of residues 165–178(180–184); L5B: Loop 5B of residues 186–187(192–195); L5C: Loop 5C of residues 192–194(200–202); L5D: Loop 5D of residues 219–224(223–229).

Fig. 1

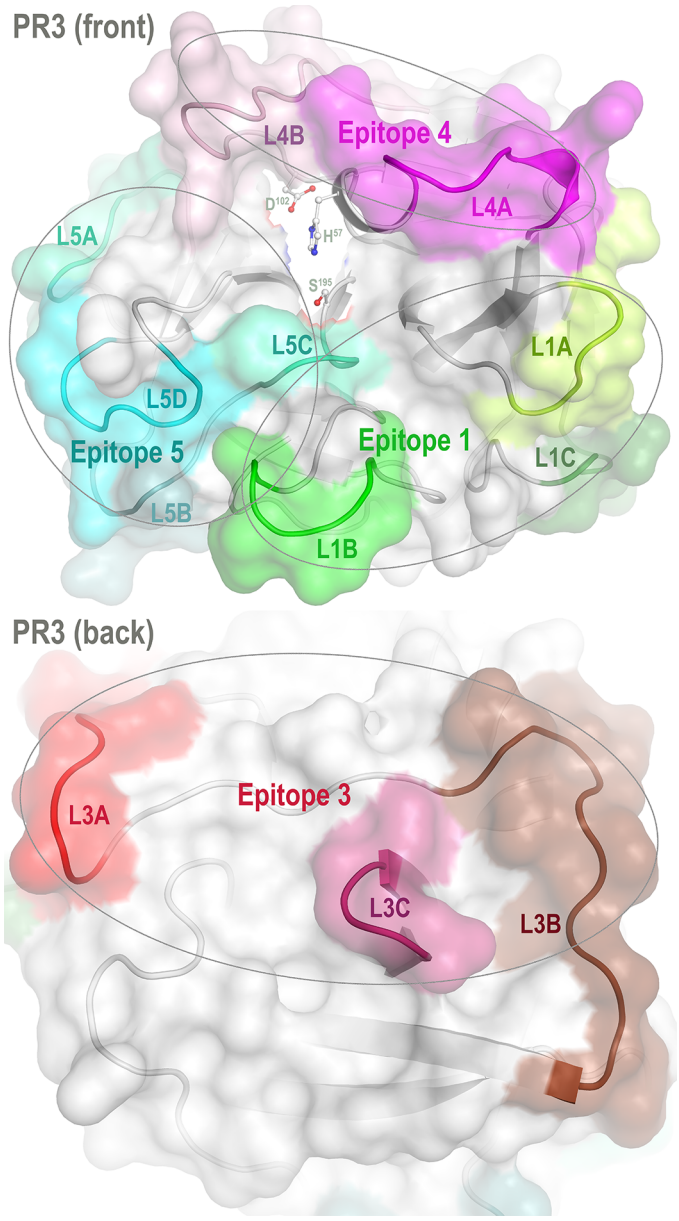
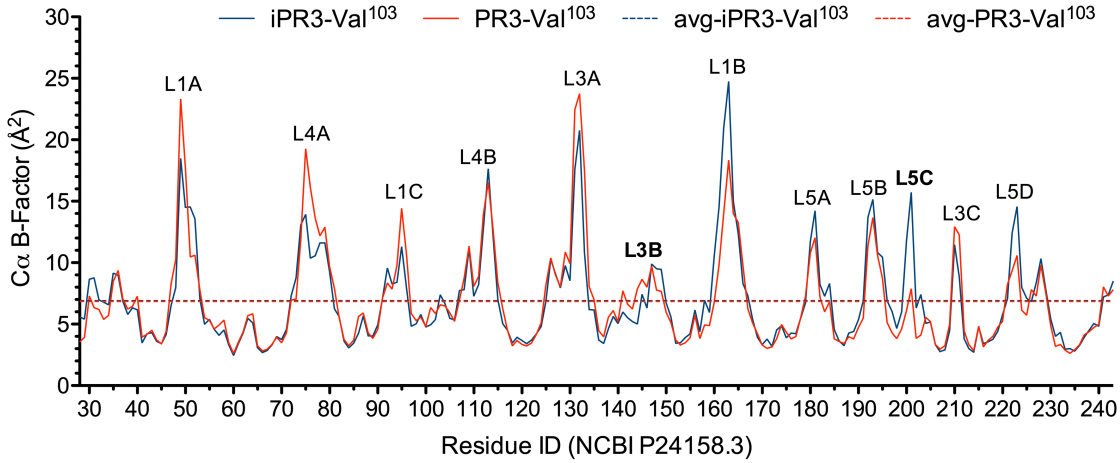

PR3 (front)
L4B
Epitope 4
L4A
D102
H57
S195
L5A
L5C
L1A
L5D
Epitope 5
Epitope 1
L1C
L5B
L1B
PR3 (back)
L3A
Epitope 3
L3C
L3B

Fig. 2


iPR3-Val103
PR3-Val103
avg-iPR3-Val103
avg-PR3-Val103
Cα B-Factor (Å2)
Residue ID (NCBI P24158.3)
L1A
L4A
L1C
L4B
L3A
L3B
L1B
L5A
L5B
L5C
L3C
L5D

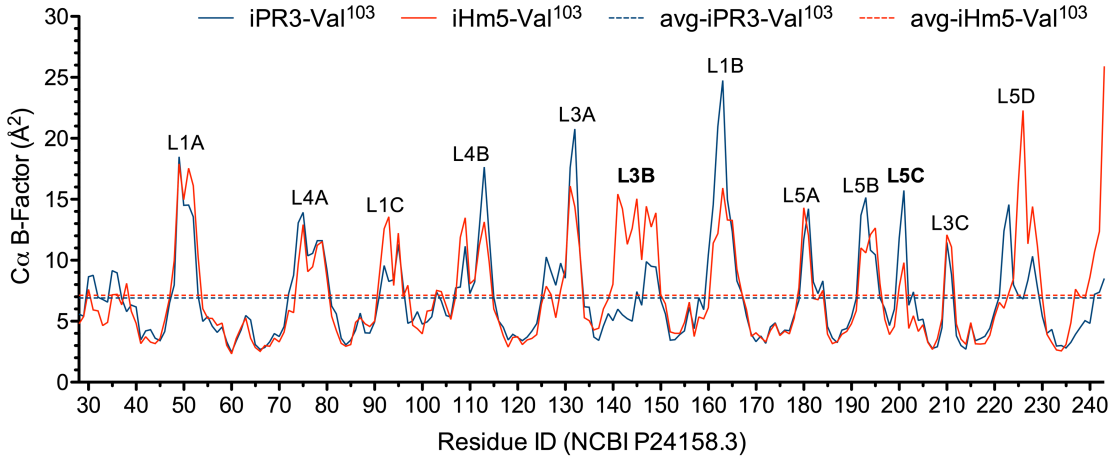

iPR3-Val103
iHm5-Val103
avg-iPR3-Val103
avg-iHm5-Val103
Cα B-Factor (Å2)
Residue ID (NCBI P24158.3)
L1A
L4A
L1C
L4B
L3A
L3B
L1B
L5A
L5B
L5C
L3C
L5D

Table S1. Alpha carbon root mean square deviations (Å) among different PR3 variants

	Xray-PR3-Ile^{103}	Comp-PR3-Ile^{103}	Comp-PR3-Val^{103}	Com-iHm5-Val^{103}
Com-PR3-Ile^{103}	1.67	0	0.62	1.63
Com-PR3-Val^{103}	1.90	0.62	0	1.56
Com-iPR3-Val^{103}	1.75	0.59	0.45	1.41
Com-iHm5-Val^{103}	2.34	1.63	1.56	0

Table S2. Alpha carbon B-factors of three PR3 variants

residue ID[a]	PR3-Val^{119}		iPR3-Val^{119}		iHm5-Val^{119}	
	mean (n = 20)	SEM[b]	mean (n = 20)	SEM[b]	mean (n = 20)	SEM[b]
28	3.56	0.31	5.60	0.44	4.79	0.31
29	3.93	0.31	5.40	0.31	5.39	0.42
30	7.28	0.80	8.64	0.58	7.62	0.49
31	6.34	0.39	8.77	1.01	5.92	0.38
32	6.20	0.41	6.97	0.45	5.84	0.34
33	5.39	0.24	6.79	0.64	4.65	0.28
34	5.72	0.29	6.57	0.53	4.91	0.25
35	8.46	0.85	9.14	1.03	7.17	0.56
36	9.37	0.83	8.96	0.46	7.22	0.49
37	7.06	0.67	6.86	0.32	6.35	0.33
38	6.24	0.43	5.79	0.63	8.11	0.47
39	6.44	0.44	6.33	0.44	5.83	0.41
40	7.26	0.84	6.17	0.73	4.83	0.31
41	3.94	0.18	3.46	0.15	3.17	0.15
42	4.17	0.27	4.21	0.37	3.73	0.13
43	4.52	0.23	4.30	0.22	3.29	0.12
44	3.72	0.17	3.62	0.13	3.16	0.18
45	3.39	0.13	3.40	0.14	3.59	0.19
46	4.41	0.22	4.16	0.16	5.08	0.29
47	8.26	0.57	6.51	0.27	6.90	0.38
48	10.26	0.67	7.96	0.40	9.97	0.49
49	23.33	1.73	18.49	1.40	17.91	1.18
50	17.78	1.86	14.51	0.97	14.97	1.03
51	10.48	1.01	14.52	1.23	17.56	2.29
52	10.62	0.67	13.58	1.51	16.14	1.72
53	7.79	0.46	7.08	0.46	10.28	0.87
54	5.53	0.32	4.99	0.19	6.03	0.27
55	5.30	0.31	5.39	0.30	5.25	0.34
56	4.61	0.26	4.59	0.30	5.24	0.35
57	4.93	0.37	4.10	0.16	4.67	0.30
58	5.33	0.22	4.54	0.30	4.84	0.22
59	3.50	0.14	3.33	0.11	3.04	0.14
60	2.61	0.10	2.45	0.07	2.34	0.08
61	3.52	0.15	3.39	0.09	3.60	0.15
62	4.36	0.18	4.29	0.15	4.49	0.19
63	5.70	0.31	5.46	0.15	5.26	0.26
64	5.85	0.43	5.12	0.22	3.62	0.15
65	3.25	0.10	3.09	0.08	2.86	0.09
66	2.83	0.08	2.68	0.06	2.54	0.07
67	2.94	0.11	2.86	0.07	3.02	0.13

68	3.36	0.13	3.31	0.13	2.94	0.09
69	3.92	0.20	3.99	0.22	3.62	0.14
70	3.49	0.13	3.76	0.20	3.31	0.12
71	4.29	0.17	4.57	0.25	4.12	0.17
72	6.95	0.40	7.28	0.60	5.89	0.25
73	7.06	0.46	8.78	0.88	5.71	0.32
74	11.84	1.54	13.07	1.37	9.68	0.71
75	19.26	3.66	13.92	1.17	12.93	1.61
76	16.13	2.05	10.37	0.83	9.06	0.67
77	13.62	0.92	10.58	0.68	9.46	0.52
78	12.19	1.25	11.60	1.10	11.22	0.81
79	12.90	1.20	11.60	1.30	11.54	0.90
80	9.70	0.49	9.29	0.86	8.66	0.64
81	7.88	0.44	6.25	0.27	5.40	0.29
82	5.56	0.31	5.62	0.29	4.25	0.22
83	3.76	0.12	3.61	0.11	3.18	0.13
84	3.26	0.15	3.07	0.10	2.95	0.15
85	3.69	0.26	3.44	0.16	3.09	0.16
86	5.63	0.58	4.26	0.24	4.90	0.25
87	5.91	0.76	5.67	0.98	5.28	0.40
88	4.29	0.26	4.06	0.17	4.78	0.20
89	3.86	0.15	4.03	0.15	4.57	0.23
90	4.58	0.22	4.95	0.28	5.02	0.27
91	7.12	0.41	7.07	0.47	8.71	0.70
92	8.36	0.47	9.57	0.72	12.59	0.85
93	7.85	0.55	8.27	0.50	13.56	1.13
94	9.76	0.68	8.41	0.46	7.92	0.54
95	14.44	1.18	11.30	0.57	12.23	1.80
96	10.63	1.08	7.91	0.33	7.02	0.63
97	5.87	0.48	4.82	0.17	7.95	1.01
98	5.29	0.32	5.04	0.30	4.68	0.31
99	5.57	0.27	5.78	0.32	4.39	0.25
100	4.80	0.20	4.76	0.20	3.96	0.14
101	6.34	0.30	4.95	0.19	5.82	0.29
102	5.86	0.26	5.37	0.18	5.91	0.24
103	6.57	0.24	7.39	0.49	7.53	0.35
104	6.49	0.42	6.67	0.39	7.42	0.34
105	5.91	0.33	5.47	0.23	6.26	0.33
106	5.23	0.22	5.35	0.25	5.16	0.26
107	6.98	0.44	7.73	0.59	7.05	0.67
108	8.72	0.59	7.81	0.71	11.90	1.93
109	11.35	0.92	11.15	0.94	13.49	1.65
110	8.06	0.56	7.27	0.48	8.06	0.48
111	8.89	0.68	8.23	0.55	8.43	0.56
112	13.91	1.28	12.71	0.76	11.24	0.80
113	16.56	1.98	17.66	1.51	13.15	1.30
114	12.60	1.61	12.03	0.94	9.89	0.91
115	8.34	0.71	6.99	0.48	6.01	0.54
116	6.25	0.41	5.03	0.22	5.02	0.30
117	4.50	0.18	4.52	0.20	3.82	0.23
118	3.23	0.11	3.47	0.18	2.89	0.09
119	3.66	0.14	3.93	0.22	3.67	0.14
120	3.36	0.10	3.67	0.16	3.73	0.13
121	3.22	0.09	3.41	0.10	3.09	0.09
122	3.47	0.13	3.70	0.11	3.46	0.13
123	4.15	0.14	4.20	0.12	3.59	0.13

124	5.12	0.15	4.80	0.19	3.90	0.18
125	8.16	0.50	6.83	0.39	6.39	0.22
126	10.38	0.70	10.26	0.62	7.87	0.49
127	9.01	0.69	9.11	0.45	7.21	0.35
128	8.03	0.47	7.94	0.52	5.29	0.33
129	10.87	1.10	9.77	0.91	7.57	1.58
130	9.92	0.95	8.53	0.37	9.28	0.75
131	22.48	2.73	17.61	1.38	16.09	2.31
132	23.75	2.39	20.77	2.37	14.38	2.82
133	17.51	2.06	10.80	0.99	10.93	1.03
134	8.05	0.71	6.19	0.40	5.28	0.36
135	6.99	0.60	6.16	0.27	5.05	0.44
136	4.47	0.23	3.70	0.13	4.28	0.45
137	3.97	0.14	3.43	0.13	4.43	0.46
138	5.55	0.28	4.61	0.19	6.01	0.67
139	6.11	0.25	5.63	0.22	6.82	0.50
140	5.20	0.10	5.05	0.13	8.03	0.39
141	7.71	0.44	5.99	0.21	15.44	1.57
142	6.64	0.49	5.53	0.30	14.24	2.08
143	6.23	0.34	5.21	0.26	11.31	1.93
144	7.86	0.56	5.00	0.26	12.60	1.62
145	8.66	0.62	7.42	0.36	15.05	1.19
146	8.00	0.46	6.31	0.38	10.03	0.53
147	9.72	0.41	9.90	0.52	14.43	1.09
148	7.77	0.27	9.51	0.56	12.73	1.15
149	7.67	0.40	9.45	0.46	13.90	1.87
150	5.96	0.28	6.73	0.38	7.17	0.53
151	5.16	0.23	5.62	0.36	6.43	0.31
152	3.70	0.10	3.45	0.09	4.13	0.13
153	3.31	0.14	3.49	0.10	4.02	0.13
154	3.49	0.09	3.90	0.11	4.02	0.10
155	3.88	0.10	4.21	0.20	4.89	0.17
156	5.73	0.52	6.15	0.47	6.57	0.47
157	3.84	0.31	4.39	0.31	3.76	0.29
158	4.92	0.29	6.94	0.57	5.34	0.65
159	4.89	0.25	5.95	0.44	5.21	0.37
160	6.75	0.38	10.46	1.45	6.11	0.34
161	9.83	0.70	14.55	2.18	11.41	1.02
162	14.05	1.29	20.99	2.59	12.20	1.01
163	18.35	1.74	24.75	2.29	15.93	1.11
164	13.99	1.82	14.99	1.15	13.31	0.95
165	13.27	1.31	12.36	0.68	13.28	1.33
166	9.76	1.20	8.31	0.49	9.26	0.87
167	6.25	0.79	7.30	0.47	7.33	0.43
168	5.12	0.35	5.42	0.32	5.96	0.32
169	4.23	0.19	3.93	0.12	3.78	0.15
170	3.32	0.14	3.34	0.10	4.06	0.27
171	3.03	0.13	3.79	0.50	3.66	0.24
172	3.14	0.15	3.18	0.20	3.31	0.12
173	3.82	0.19	4.54	0.18	4.24	0.24
174	4.96	0.29	4.87	0.21	4.80	0.23
175	4.34	0.28	3.88	0.26	3.83	0.18
176	3.81	0.14	4.26	0.25	4.19	0.18
177	3.99	0.18	4.24	0.26	3.96	0.13
178	5.56	0.24	5.50	0.29	5.26	0.25
179	7.24	0.66	6.72	0.41	7.75	0.61

180	10.82	1.03	11.74	0.69	14.32	1.44
181	12.04	0.88	14.22	1.21	12.05	0.87
182	7.22	0.62	8.24	0.49	6.87	0.44
183	6.00	0.29	7.27	0.50	6.76	0.56
184	6.81	0.42	8.31	0.56	7.53	0.53
185	3.81	0.18	4.57	0.36	3.94	0.23
186	3.60	0.17	3.63	0.13	3.16	0.19
187	3.42	0.10	3.26	0.13	3.33	0.13
188	3.96	0.26	4.26	0.29	3.93	0.16
189	3.81	0.21	4.40	0.38	4.17	0.19
190	4.72	0.33	5.32	0.79	4.78	0.18
191	5.48	0.31	6.89	0.84	5.86	0.38
192	11.37	0.89	13.69	2.05	11.01	1.22
193	13.67	1.31	15.15	2.30	10.60	0.93
194	10.60	0.61	10.83	1.01	12.10	1.10
195	8.73	0.49	10.45	0.87	12.66	1.31
196	5.13	0.27	6.91	0.63	8.07	0.66
197	4.28	0.14	6.04	0.56	5.16	0.26
198	3.83	0.18	4.65	0.41	3.91	0.13
199	4.62	0.25	5.99	0.37	4.58	0.21
200	6.13	0.52	11.85	1.40	7.84	0.47
201	7.88	0.56	15.72	1.54	9.81	1.02
202	3.86	0.28	6.30	0.57	4.41	0.36
203	4.11	0.22	7.41	1.22	5.45	0.36
204	5.56	0.52	5.07	0.28	4.18	0.20
205	5.21	0.30	5.17	0.36	4.71	0.20
206	3.35	0.11	3.33	0.14	3.27	0.09
207	2.97	0.10	2.78	0.09	2.69	0.07
208	3.26	0.08	2.91	0.09	3.55	0.16
209	4.98	0.18	4.47	0.14	5.22	0.28
210	12.93	0.64	11.47	0.78	12.09	0.96
211	12.26	0.63	8.85	0.41	11.09	0.57
212	4.44	0.14	3.81	0.12	4.78	0.22
213	3.34	0.13	3.02	0.06	3.54	0.21
214	2.85	0.09	2.70	0.06	3.11	0.17
215	4.83	0.23	4.78	0.25	4.89	0.25
216	3.15	0.22	3.40	0.19	3.14	0.11
217	3.69	0.40	3.58	0.19	3.13	0.10
218	4.01	0.27	3.77	0.24	3.16	0.11
219	4.77	0.28	4.42	0.20	3.85	0.16
220	5.24	0.42	5.76	0.45	5.35	0.24
221	8.22	0.59	7.06	0.47	6.56	0.48
222	9.35	0.81	12.41	1.09	6.07	0.38
223	10.59	0.80	14.58	1.43	7.32	0.40
224	6.13	0.33	7.99	0.58	8.49	0.46
225	5.73	0.25	7.11	0.50	16.39	0.92
226	7.81	0.43	6.82	0.67	22.29	3.15
227	7.29	0.36	8.37	0.93	11.36	0.83
228	9.81	0.53	10.33	1.31	14.41	2.07
229	7.50	0.40	7.78	0.75	11.17	1.10
230	4.84	0.24	5.40	0.67	6.98	0.51
231	3.18	0.17	4.01	0.67	3.93	0.19
232	3.35	0.14	4.33	0.64	3.24	0.11
233	2.88	0.12	2.96	0.10	2.65	0.08
234	2.63	0.10	3.00	0.14	2.58	0.07
235	2.90	0.10	2.82	0.09	3.10	0.10

236	3.33	0.10	3.27	0.12	4.87	0.30
237	4.13	0.15	3.93	0.16	7.64	0.65
238	4.38	0.16	4.51	0.18	6.99	0.49
239	4.74	0.22	5.06	0.21	6.91	0.53
240	4.98	0.30	4.83	0.17	8.61	0.74
241	8.03	0.65	7.22	0.37	10.68	0.73
242	7.29	0.60	7.39	0.31	12.37	1.07
243	7.81	0.68	8.50	0.64	25.91	2.54

[a]The residue numbering here is identical to that of the human PR3 sequence (NCBI P24158.3). [b]SEM: Standard error of the mean.

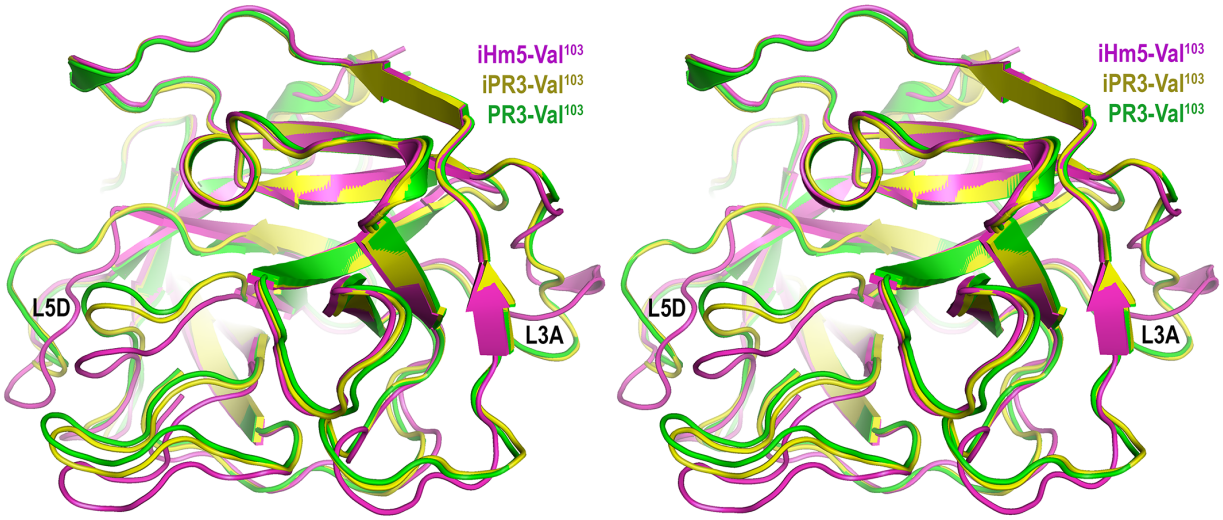



Fig. S1. Superimposed computer models of three PR3 variants. The backbone conformations in cross-eye stereo view show their close similarity and local differences in two loops. These loops are labeled as L3A for residues 110–117 and L5D for residues 219–224.