Supporting information for:

Computational Modeling of Human Paraoxonase 1: Preparation of Protein Models, Binding Studies, and Mechanistic Insights

Toby T. Sanan[a], Sivaramakrishnan Muthukrishnan[a], Jeremy M. Beck[a], Peng Tao[a], Carrigan J. Hayes[a], Tamara C. Otto[b], Douglas M. Cerasoli[b], David E. Lenz[b], and Christopher M. Hadad[a]*

[a] Department of Chemistry, 100 West 18th Avenue, Ohio State University, Columbus, Ohio 43210

[b] 3100 Ricketts Point Rd, Physiology and Immunology Branch, Research Division, US Army Medical Research Institute of Chemical Defense, Aberdeen Proving Ground, Maryland 21010

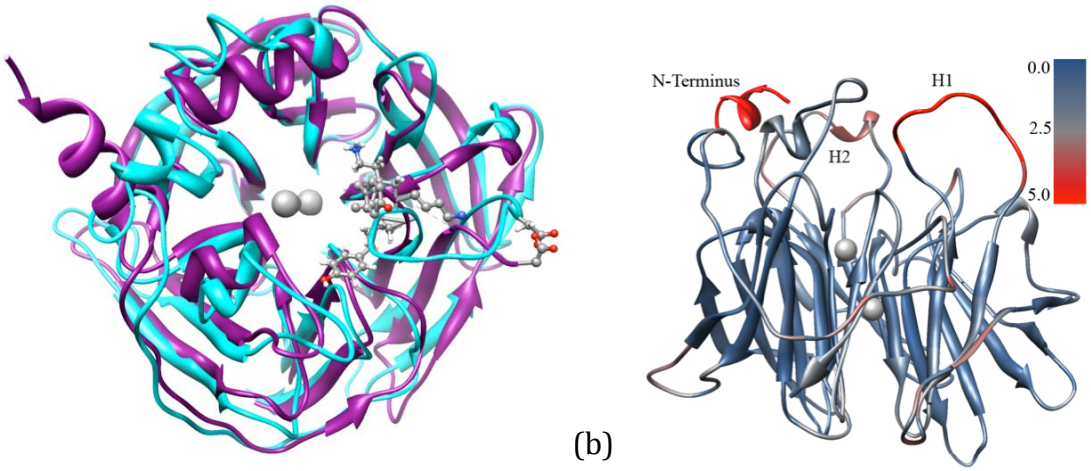
* Corresponding author: hadad.1@osu.edu

Preparation of the G2E6 Model

Treatment of residues 72-79

Residues 72 to 79 were unresolved in the crystal structure of G2E6, and due to proximity to the active site, reconstruction of this loop was desirable for the protein model. The structure of the related enzyme diisopropylfluorophosphatase (DFPase) was previously employed for generation of a homology model of huPON1 WT (pdb ID: 1XHR), and in this structural model the loop containing residues 72 to 79 was intact. We used this loop as a template for reconstruction of the loop in our protein model system. A comparison of the loop in the G2E6 model and the crystal structure is shown in Figure S1. There is some rotation of the loop backbone, which alters the positioning of residues L69, Y71, and particularly K70 in proximity to the active site, but the overall motion of the anchor points of the loop into the β-propeller is minor. The high flexibility of these residues, as indicated by the inability to resolve this portion of the protein in the X-ray crystal structure, is reflected in their movement in the MD simulations. Similar findings were reported in a recent publication of a model of huPON1,

showing considerable changes in the conformation of these residues in both the apo-protein and in models with substrate bound.[1]



(a) (b)

Figure S1. (a) Comparison of the G2E6 computational model (teal) and the X-ray crystal structure (pdb ID: 1V04, magenta). Residues L69, K70, Y71, and D80 are shown in ball and stick mode, and the calcium ions are in gray. (b) G2E6 computational model colorized based on backbone RMSD changes relative to the crystal structure; the color on the H1 (72-79) loop is arbitrarily set to the maximum value, as the residues in the crystal structure are missing.

Interactions of K192

The K192 loop is quite flexible, and over the course of MD showed significant lateral movement on top of the protein, such that the interactions between K192 and S166/D183 were somewhat time-dependent; when the loop was oriented toward the latter residues, hydrogen bonds were formed to both residues, but movement away inhibited these interactions. It can be seen from the RMSD trace (Figure S2) that the movement of the loop influenced the interactions both to K192 and between S166 and D183.

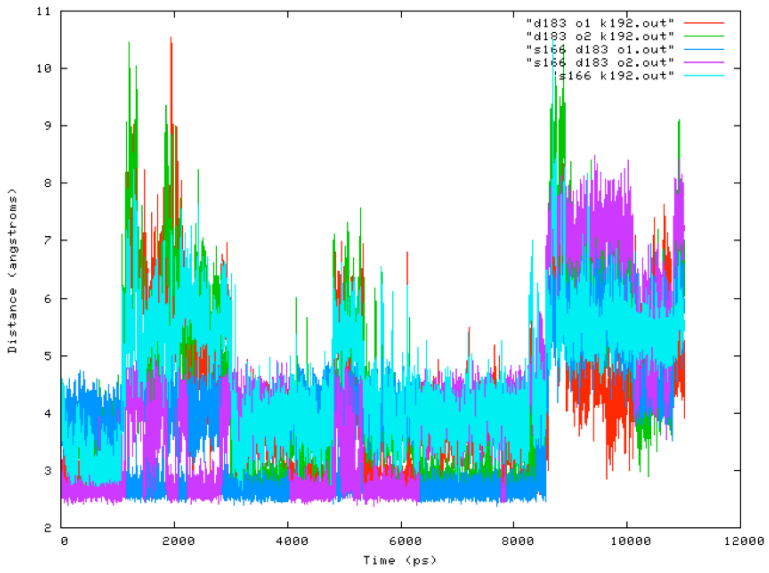


Figure S2. Overlay of distance measurements for D183–K192, S166–D183, and S166–K192 monitored over the terminal 11 ns of MD in the G2E6 protein model.

As the distance plots indicate, the distances between the three residues were quite stable over large portions of the simulation, with two stable configurations; one with direct hydrogen bonding between K192 and S166/D183, the other with K192 more distant.

As a precursor to more detailed studies of the role of residue 192 in substrate binding, models were prepared containing K192Q and K192R mutations. The starting structure for these models was the 'relaxed' G2E6 protein model, and an identical minimization and heating regime was followed. A total of 5 ns of MD was performed on these models, after which analysis of the RMSD suggested the proteins had fully relaxed (Figure S3).

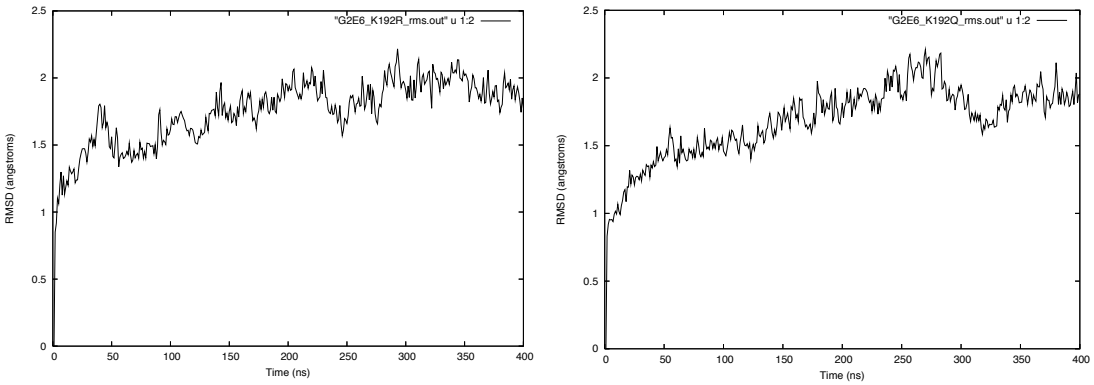


Figure S3. RMSD plots for the G2E6 K192R (left) and K192Q (right) mutants over the terminal 4 ns of MD suggest that the models had stabilized.

The structural changes in these two models suggest that these mutations can alter the potential electrostatic and hydrogen bonding interactions in this region of the active site, which may have a differential role in terms of binding energetics and orientaions for the V-series agents, specifically. More simulations are in progress to model the S166N/K192R and S166N/K192Q double mutants, to more closely represent the WT protein.

Analysis of MD simulation stability

Over the course of the MD simulations, the stability of the G2E6 protein model was monitored by evaluation of the total, potential, and kinetic energy over time, as well as other parameters. This can be seen in Figure S4. These clearly illustrate the rapid equilibration of the protein model, and suggest that the simulations did not develop any instabilities.

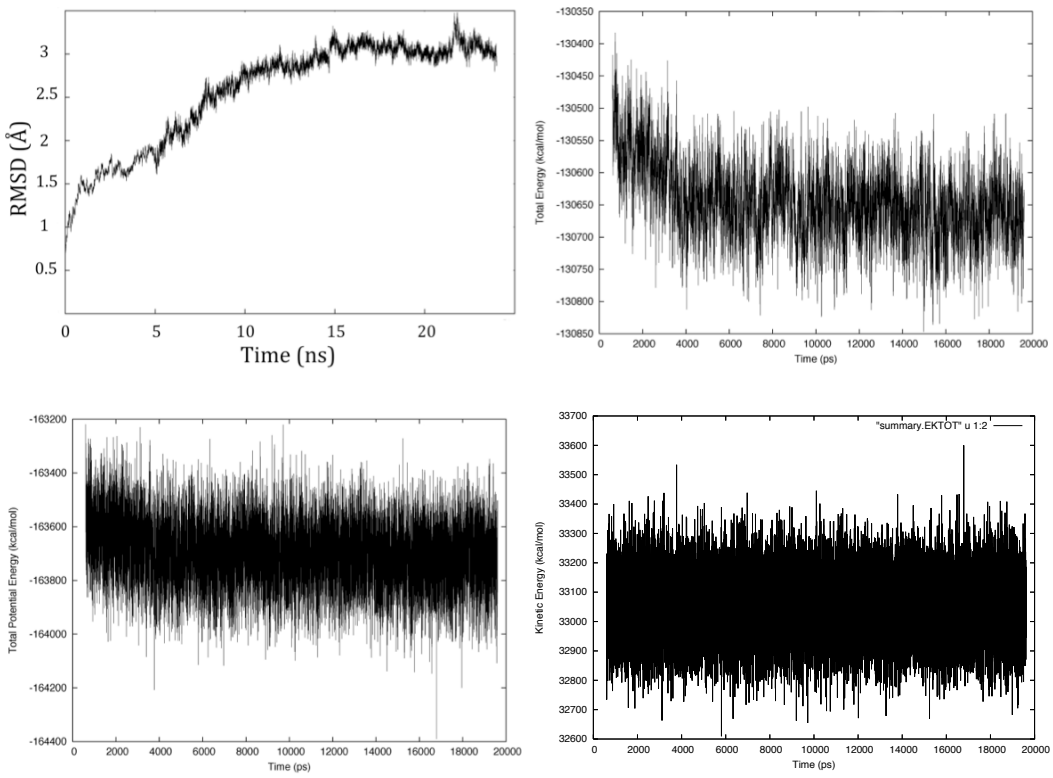


Figure S4. RMS, total energy, total kinetic, and total potential energy modeled over the course of over 20 ns of MD simulations for the G2E6 protein model, showing the stability of the protein model under the simulation conditions.

Autodock 4.0 Preparation and Parameters

For the docking simulations, the program Autodock Tools 1.5.2[2] was employed for receptor model preparation. As described in the text, a series of 8 snapshots from the terminal 4 ns of MD simulations were extracted from the G2E6 protein model system, and were imported into Autodock using their AMBER FF03 atomic charges. Non-polar hydrogens and their charges were merged into adjacent heavy atoms prior to generation of the scoring grid. Default torsion parameters were employed for the six flexible residues, K70, H115, F222, I291, F292, and V346 (Figure S5). Residue K70 in particular was quite altered in position in the G2E6 model, and by including it as a flexible residue, we hoped to remove any possible perturbations of the active site arising from its mobility. Analysis of the docked modes showed no significant change in

binding modes between the different snapshots. Representative binding modes for the various substrate classes are shown below, along with docking energy scores.

For generation of the Autodock scoring grid, an 18.75 x 15.00 x 15.00 $Å^3$ box was generated centered ~3 Å above the 'catalytic' calcium ion in the direction of the putative active site. A 0.375 Å grid spacing was employed with a distance-dependent dielectric constant.

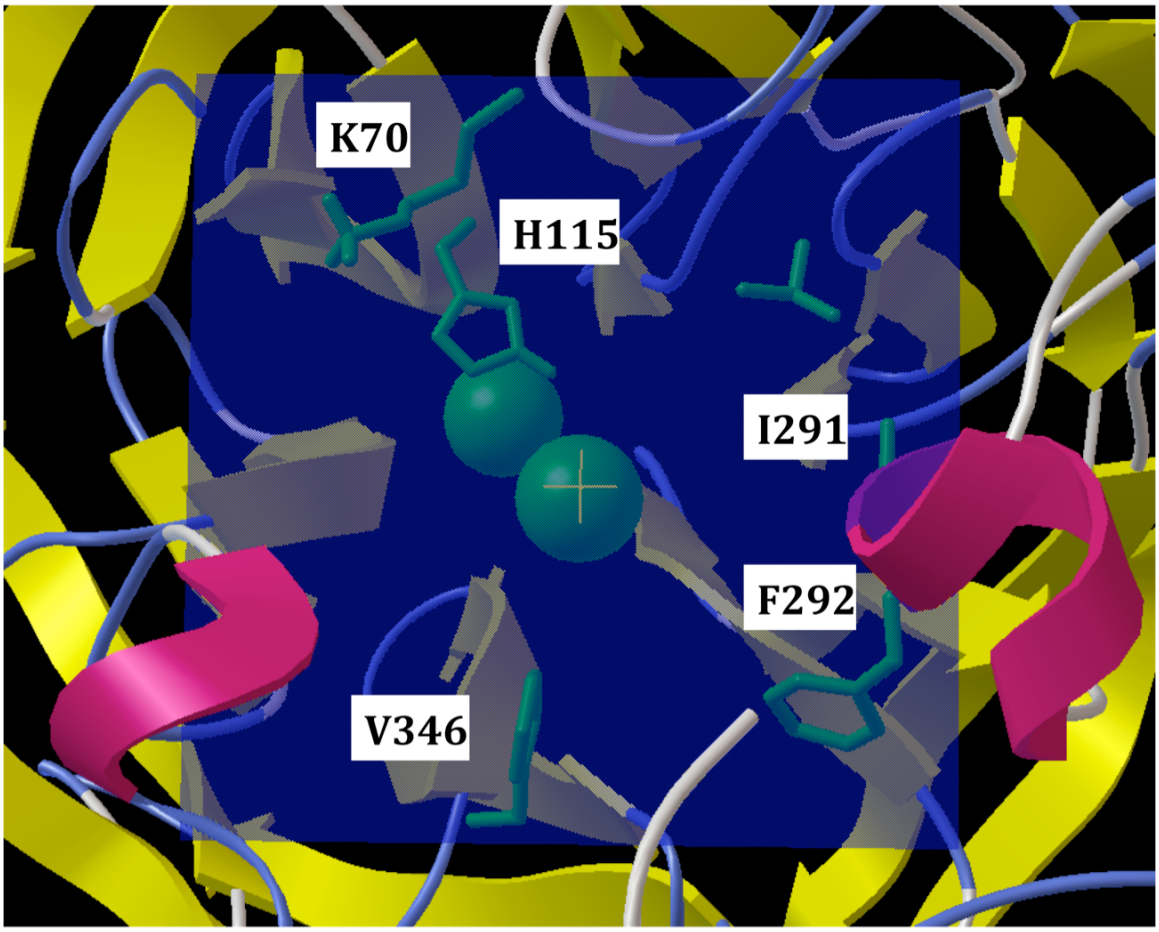


Figure S5. The 18.75 x 15.00 x 15.00 $Å^3$ scoring grid employed in the autodock simulations. The box was centered ~3 Å above the 'catalytic' calcium ion. Flexible residues are shown in green, and included K70, H115, F222, I291, F292, and V346.

To obtain charges for the ligands (Figure S6), each structure was optimized at the B3LYP/6-31+G(d,p) level of theory, after which ChelpG[3] nuclear-centered atomic charges were generated at the B3LYP/6-311+G(d,p)//B3LYP/6-31+G(d,p) level of theory using the program

Gaussian03.[4,5,6,7] For ligand preparation, Autodock Tools default torsional definitions were employed to allow for ligand conformational sampling. Substrate poses were clustered based upon an RMS threshold of 2.0 Å. Representative poses for the G- and V-series nerve agents are shown in Figures S7-S10. Due to the size of the active site, and the autodock box employed, only some poses were actually coordinated to the calcium ion following the docking simulations. The percentage of 'bound' poses can be seen in Table S1.

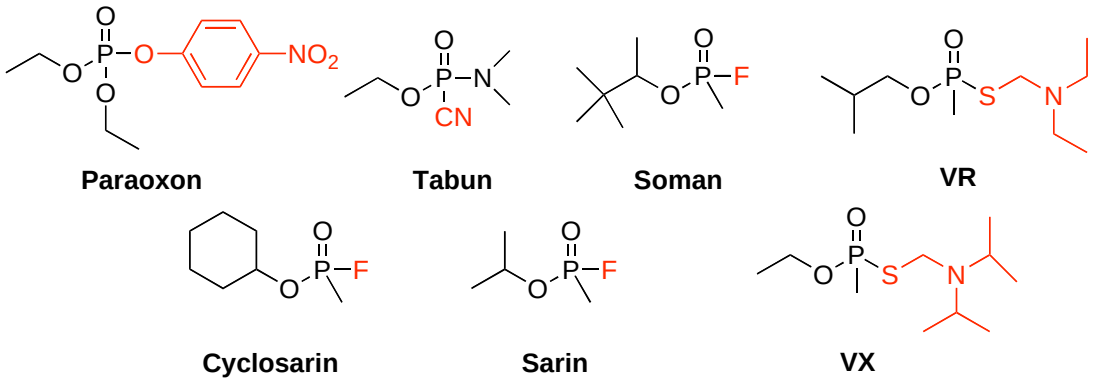
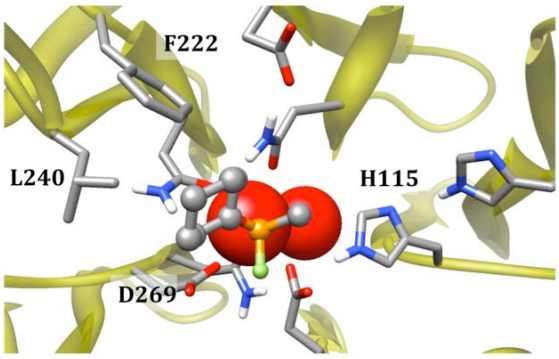
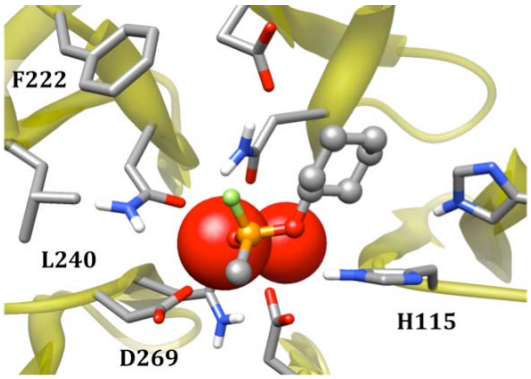


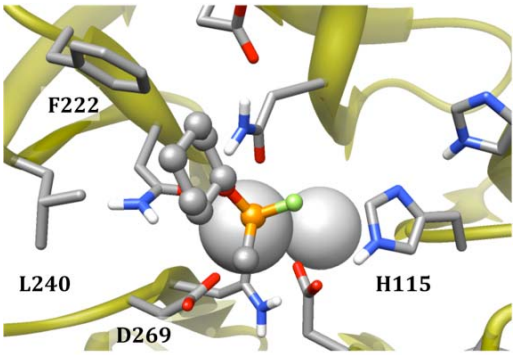
Figure S6. Organophosphorus compounds employed for binding studies on huPON1. For the V-series agents VR and VX, the question of protonation of the amino group was considered, and was found to have considerable implications for substrate interactions with the protein. Accordingly, both the neutral and protonated species were studied.



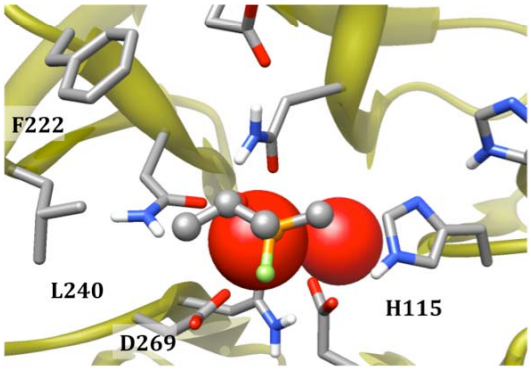
P_R Cyclosarin Docking Pose
Energy Score: -11.7 kcal/mol



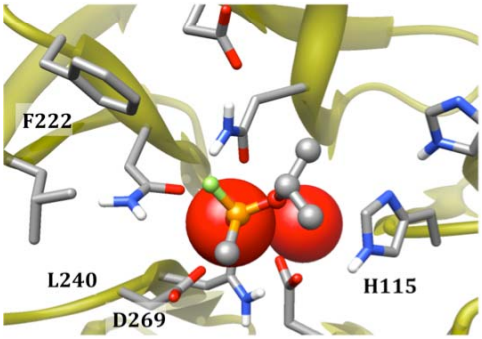
P_R Cyclosarin Docking Pose
Energy Score: -10.8 kcal/mol



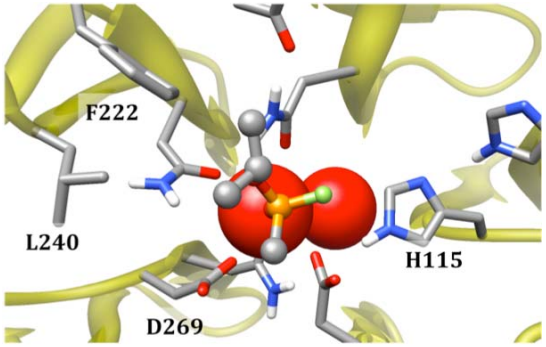
P_S Cyclosarin Docking Pose
Energy Score: -12.3 kcal/mol



P_R Sarin Docking Pose
Energy Score: -11.9 kcal/mol



P_R Sarin Docking Pose
Energy Score: -10.7 kcal/mol



P_S Sarin Docking Pose
Energy Score: -11.6 kcal/mol

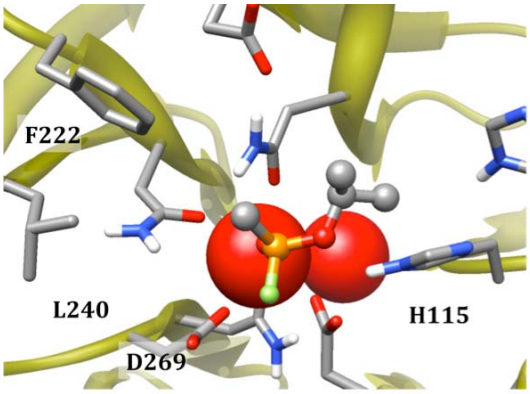


Figure S7. Docking poses obtained for sarin and cyclosarin in the G2E6 model. The docking poses demonstrate a bias for orientation of the leaving group fluoride *anti* to D269 for the P_S enantiomers, with the *i*-Pr and cyclohexyl groups preferentially placed in the L240/F222 and H115/H134 pockets; in the former orientation the P_S isomers place the fluoride *anti*, while the P_R isomers orient it toward D269.

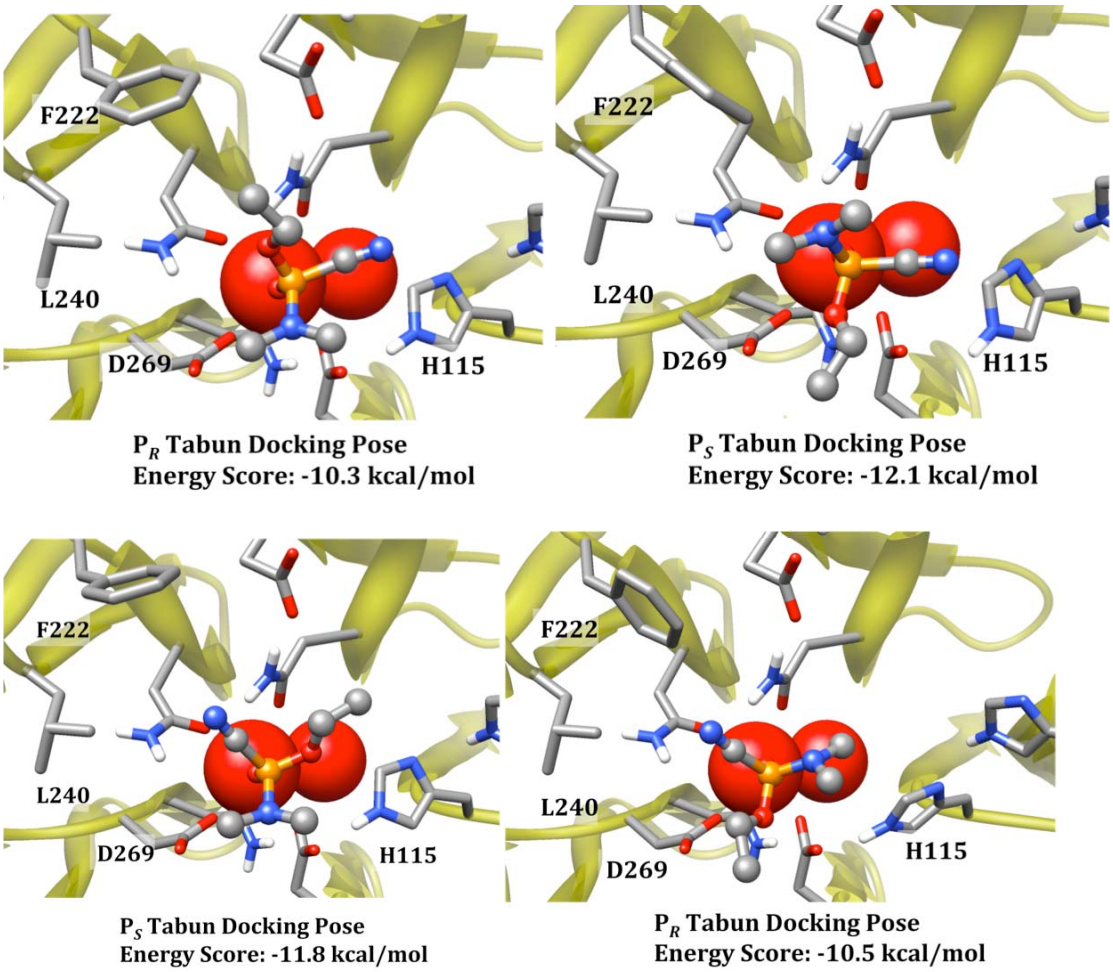
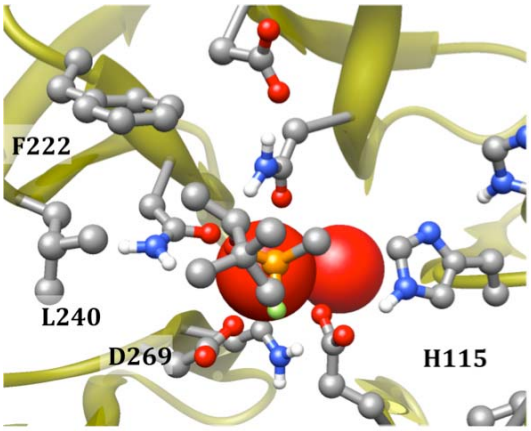
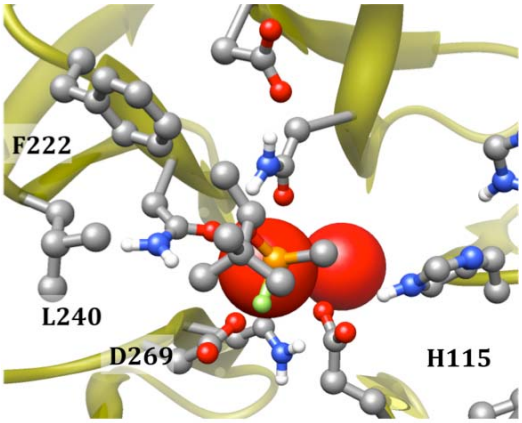


Figure S8. Docking poses obtained for tabun in the G2E6 model. For tabun the peripheral groups on phosphorus are all similar in bulk, and accordingly little preference was observed in terms of their orientation in the active site.

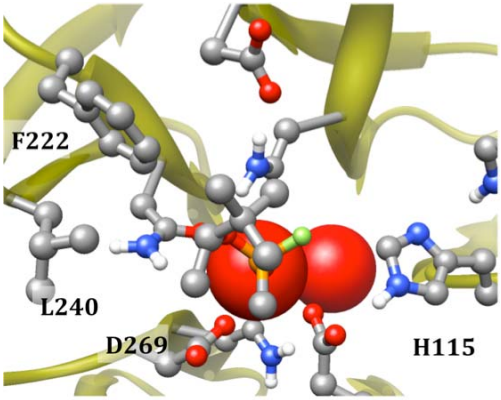


Soman P- C- Docking Pose
Energy Score: -10.4 kcal/mol

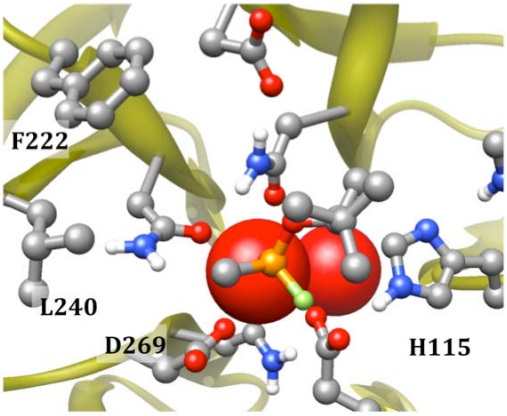


Soman P- C- Docking Pose
Energy Score: -10.2 kcal/mol

(continued on next page)



Soman P+ C- Docking Pose
Energy Score: -9.0 kcal/mol



Soman P+ C- Docking Pose
Energy Score: -8.3 kcal/mol

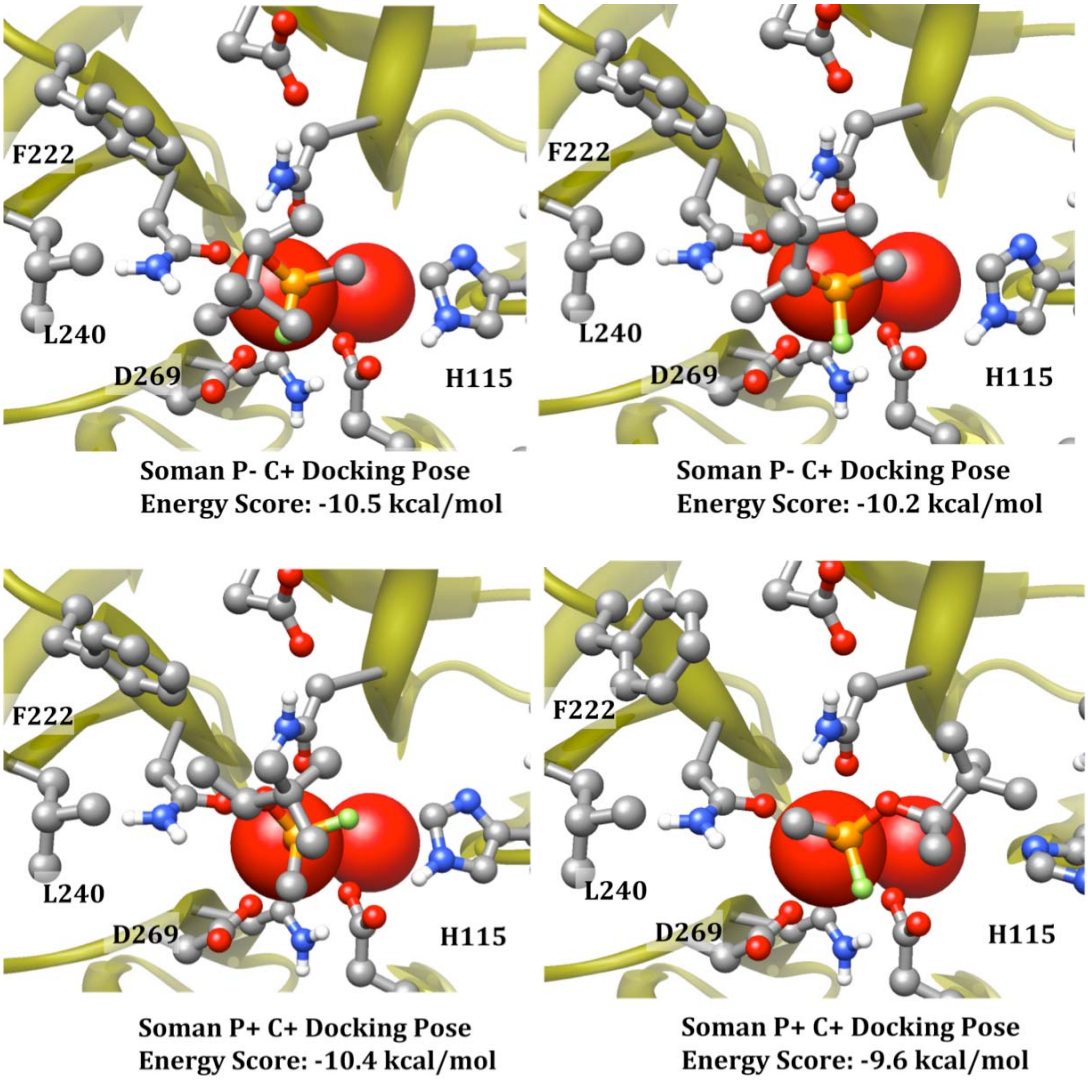


Figure S9. Docking poses obtained for the soman diastereomers in the G2E6 model. Experimentally, the rates of hydrolysis are $P^{+}C^{+}$ >> $P^{+}C^{-}$ > $P^{-}C^{+}$ >> $P^{-}C^{-}$.[8] In the docking modes the diastereomers with P^{+} are able to orient the fluorine *anti* to D269, while the P^{-} diastereomers are unable to do so due to the steric constraints of the bulky alkyl group, which is able to orient primarily in the pockets proximal to L240/F222, and H115/H134.

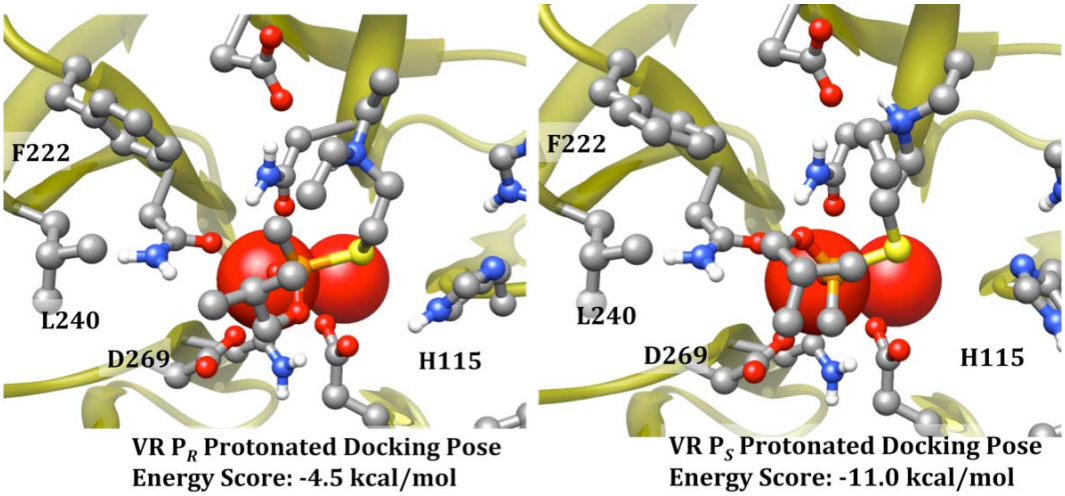


Figure S10. Docking poses obtained for the V-series nerve agents.

Table S1.

Agent:	Percent of Poses within 2 Å from Calcium:
GB (P_R)	98
GB (P_S)	92
GD C^+P^+	69
GD C^+P^-	71
GD C^-P^+	76
GD C^-P^-	81
GF^+	92
GF^-	89
Paraoxon	25
Tabun (P_R)	52
Tabun (P_S)	74
VR (P_R)	12
VR (P_S)	7
VX (P_R)	31
VX (P_S)	39

Estimations of Free Energy of Binding Using Molecular Dynamics

To obtain optimized, bound conformations of nerve agents, we started with the docking orientations of OPs in the G2E6 model system. The proper protonation states for the OPs were restored prior to MD, and the GAFF force field was used for additional force field terms. To ensure the proper treatment of ligand-receptor interactions, a moderate 20 kcal/mol–$Å^2$ restraint

was employed for the calcium-phosphoryl oxygen bond coordinate, with the form of the surface a flat-bottomed parabola, as is default in restrained bond distances in AMBER 9.0. The restraint had a value of zero between 2.5 and 4.0 Å, and increased parabolically between 2.0 and 2.5, and 4.0 to 5.0 angstroms, by the equations

$$k_2 (R - r_2)^2 \quad (1)$$

$$k_3 (R - r_3)^2 \quad (2)$$

where r_2 and r_3 are 2.5 and 4.0 Å, respectively, R is the actual bond distance, and k_2 and k_3 are 20 kcal/mol–Å^2. Outside the boundaries, the energy of the restrained bond increases linearly with a slope equal to the ends of the parabolic restraints.

Following 1 ns of restrained MD simulations, the restraint was removed, and unrestrained simulations were performed for another 1 ns. The purpose of the restraint was to restrict mobility of the coordinated OPs prior to relaxation of the protein–ligand contacts, as some simulations had demonstrated some instability in the coordination of substrate to calcium, which we attributed to inadequate relaxation time for ligand-receptor contracts during the minimization and heating steps. This was successful for some simulations; however, for the P_S enantiomers of VR and V5 stable, bound conformations in an orientation with the leaving group *anti* to D269 were not identified in the same simulation series.

The question of the stability of calcium coordination following MD was also examined, and while some movement of the ‘catalytic’ calcium was observed, in general the coordinating residues had minor side-chain movements, but the backbone motions were minor, and while systematic mapping of potential changes in coordination was not attempted, the aggregated MD simulations of ligand-bound structures did not identify any systematic changes.

Energies obtained for the V-series agents following MM-PBSA and MM-GBSA energy decompositions are shown in Table S2. For the P_R agents, poses with the leaving group *anti* to D269 and oriented into the D269 cavity were both obtained, allowing for an energetic comparison of the two. For VX, the pose with the leaving group *anti* was lower in energy, although in some cases, this was within the margin of error. For VR, the energies were comparable. For the P_S agents, the binding motif with the thioalkylamino fragment in the D269 pocket was not observed in the MD simulations, similar to the docking simulations above. This provides additional support for the mechanistic hypothesis, with the stereoselectivity of huPON1 arising due to orientational preferences in the active site; for the P_R agents placement of the leaving group *anti* to D269 requires placement of the bulkier *O*-alkyl fragment into the more congested space in the active site, whereas the P_S agents can orient this group into the more spacious pocket around H285 and L240.

Table S2. Leaving group orientations and calculated free energies of binding for the P_R and P_S V-series nerve agents in the G2E6 model. Poses in **bold** were dissociative from coordination to calcium, despite multiple simulation attempts.

Nerve	P_R Enantiomer LG in H115 Pocket		P_S Enantiomer LG in H115 Pocket		P_R Enantiomer LG in H285 Pocket	
Agent	MM-PBSA	MM-GBSA	MM-PBSA	MM-GBSA	MM-PBSA	MM-GBSA
VR	-31.0 ± 4.2	-27.4 ± 3.0	**-16.6 ± 6.1**	**-20.6 ± 3.6**	-31.4 ± 4.2	-26.7 ± 3.4
VX	-28.8 ± 4.2	-23.4 ± 3.0	-22.3 ± 5.6	-27.7 ± 4.7	-19.1 ± 4.3	-14.5 ± 2.9

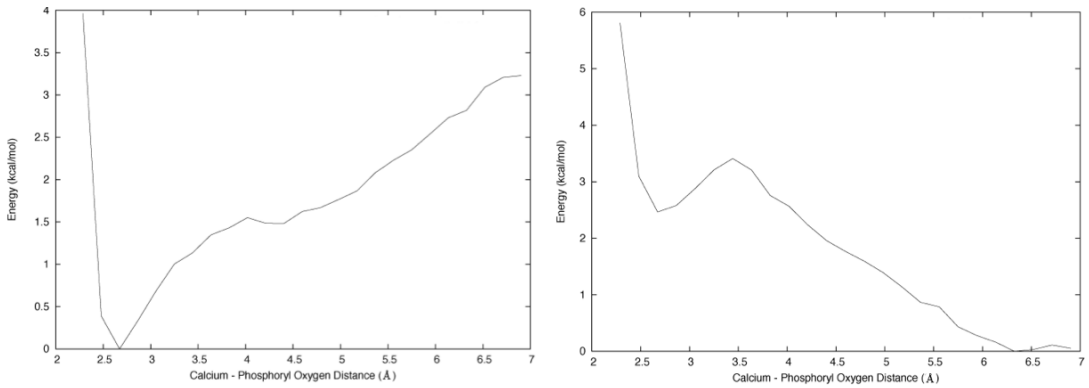


Figure S11. Potential energy surfaces for dissociation of V-series nerve agents from coordination to the calcium in the G2E6 protein model for P_S VX (left) and VR (right), of which VR was dissociative under the MD conditions.

References:

[1] X. Hu, X. Jiang, D. E. Lenz, D. M. Cerasoli, A. Wallqvist, *Proteins* **2009**, *75*, 486-498.

[2] M. F. Sanner. *J. Mol. Graphics Mod.* **1999**, *17*, 57-61.

[3] C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **1990**, *11*, 361-373

[4] (a) A. D. Becke, J. Chem. Phys. **1993**, 98, 5648-5656. (b) C. Lee, W. Yang, R. G. Parr, Phys. Rev. B **1988**, 37, 785-789.

[5] J. Labanowski, in *Density Functional Methods in Chemistry*. Ed. Springer-Verlag, Heidelberg, **1991**. (b) R. G. Parr, Y. Weitao, in *Density-Functional Theory in Atoms and Molecules*, Oxford University Press, New York, **1989**.

[6] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, T. Vreven, Jr. K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, Y. Nakajima, O. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, M. C. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, Gaussian 03, Revision B.04, Gaussian, Inc., Pittsburgh PA, 2003.

[7] (a) P. J. Stevens, F. J. Devlin, C. F. Chablowski, M. F. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623-11627. (b) A. D. Becke, *Phys. Rev. A.* **1988**, *38*, 3098-3100. (c) A. D. Becke, *J. Chem. Phys.*

1992, *96*, 2155-2160. (d) A. D. Becke, *J. Chem. Phys.* **1992**, *97*, 9173-9177. (e) A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648-5652.

[8] D. T. Yeung, J. R. Smith, R. E. Sweeny, D. E. Lenz, D. M. Cerasoli, *FEBS J.* **2007**, *274* 1183-1191.