Computational design of receptors for an organophosphate surrogate of the nerve agent soman

Malin Allert*†, Shahir S. Rizk*†, Loren L. Looger*, and Homme W. Hellinga*‡§

Departments of *Biochemistry and ‡Pharmacology and Molecular Cancer Biology, Box 3711, Duke University Medical Center, Durham, NC 27710

Edited by James A. Wells, Sunesis Pharmaceuticals, South San Francisco, CA, and approved April 13, 2004 (received for review February 24, 2004)

We report the computational design of soluble protein receptors for pinacolyl methyl phosphonic acid (PMPA), the predominant hydrolytic product of the nerve agent soman. Using recently developed computational protein design techniques, the ligand-binding pockets of two periplasmic binding proteins, glucose-binding protein and ribose-binding protein, were converted to bind PMPA instead of their cognate sugars. The designs introduce 9–12 mutations in the parent proteins. Twelve of 20 designs tested exhibited PMPA-dependent changes in emission intensity of a fluorescent reporter with affinities between 45 nM and 10 μM. The contributions to ligand binding by individual residues were determined in two designs by alanine-scanning mutagenesis, and are consistent with the molecular models. These results demonstrate that designed receptors with radically altered binding specificities and affinities that rival or exceed those of the parent proteins can be successfully predicted. The designs vary in parent scaffold, sequence diversity, and orientation of docked ligand, suggesting that the number of possible solutions to the design problem is large and degenerate. This observation has implications for the genesis of biological function by random processes. The designed receptors reported here may have utility in the development of fluorescent biosensors for monitoring nerve agents.

computational protein design | fluorescent biosensor

The most commonly used methods for manipulating ligand-binding specificity are empirical, using either the immune system to generate antibodies or directed evolution of proteins (1). These approaches lose in generality because they either are limited to a particular class of proteins (e.g., antibodies) or are constrained by selection methods, sequence diversity, or library size. Structure-based computational design methods, on the other hand, offer enormous generality for manipulating protein structure and function (2–5). However, limitations in the description of the molecular interactions (6, 7) and the immense combinatorial complexity of the sequence design problem (8) present significant barriers. Nevertheless, recent experiments have shown that, when powerful combinatorial search algorithms, accurate representation of molecular interactions, and state-of-the-art computer hardware are combined, impressive successes are obtained (9–13). Recently, we reported the radical redesign of ligand-binding sites by computational design, converting several sugar- or amino acid-binding proteins into high-affinity, specific receptors for trinitrotoluene (TNT), lactate, or serotonin (11). Here we report a further investigation into the scope of the computational (re-)design of ligand-binding sites, and present the conversion of ribose- and glucose-binding proteins into receptors for pinacolyl methyl phosphonic acid (PMPA), an organophosphate surrogate of the nerve agent soman (Fig. 1) (14).

Glucose- (GBP) and ribose-binding (RBP) proteins are members of the *Escherichia coli* periplasmic binding protein (PBP) superfamily (15). Their structures have been determined to high resolution (16, 17). Each consists of a single polypeptide chain that folds into two domains connected by a hinge region, with a ligand-binding site located in the interface between the two domains. The proteins adopt two conformations: a ligand-free open form and a ligand-bound closed form. These conformations interconvert via hinge-bending motions (18) that have been exploited to design reagentless optical or electrochemical biosensors (19, 20). Redesign of the ligand-binding specificity of PBPs (11, 21–23) is therefore a powerful route for systematic biosensor development (24).

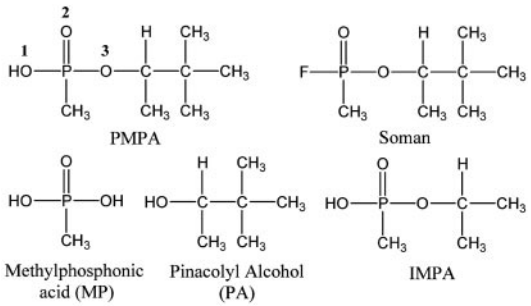

Fig. 1. Structures of PMPA, soman, and related molecules.

One application for the computational design of binding sites is the development of biosensors that detect chemical pollutants or threats. PMPA is a relatively nontoxic surrogate and the predominant hydrolytic degradation product of soman, a member of the organophosphate nerve agent family (Fig. 1) (14). Although the destruction of these agents was mandated by the 1993 Chemical Weapons Convention (25), some have been deployed in recent conflicts (26, 27) and terrorist attacks (28). Soman is a potent suicide inhibitor of acetylcholinesterase, and degrades rapidly upon exposure to water, forming PMPA (14). PMPA is only found after exposure to soman (29), and may even be present in the leading edge of a nerve agent cloud (30). Detection of PMPA is therefore important for weapons control, postincident exposure determination, and cleanup (28), and may prove useful as an attack indicator in a stand-off detector. Neither PMPA nor soman have an intrinsic chromophore or fluorophore. Current detection methods depend on acetyl cholinesterase inhibition (31), antibodies (32), capillary electrophoresis (33, 34), conductivity (35), or mass spectrometry (36). Therefore, a reagentless fluorescent biosensor for PMPA that responds rapidly and continuously is of great potential benefit for monitoring and control of this agent.

Materials and Methods

Materials. PMPA racemate was purchased from Aldrich; pinacolyl alcohol and methyl phosphonate were purchased from Fluka; isopropyl methyl phosphonic acid (IMPA) was purchased from Cerilliant (Round Rock, TX); buffers and salts were purchased from Sigma; mutagenic oligonucleotides were purchased from Operon Technologies (Alameda, CA); pET21a plasmid was purchased from Novagen; and iodoacetamidofluorescein (fluorescein,

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: PMPA, pinacolyl methylphosphonic acid; GBP, glucose-binding protein; RBP, ribose-binding protein; IMPA, isopropyl methylphosphonic acid; IAF, iodoacetamidofluorescein; PCS, primary complementary surface; MP, methylphosphonate; PA, pinacolyl alcohol; NBDE, *N*-((2-(iodoacetoxy)ethyl)-*N*-methyl)amino-7-nitrobenz-2-oxa-1,3-diazole.

†M.A. and S.S.R. contributed equally to this work.

§To whom correspondence should be addressed. E-mail: hwh@biochem.duke.edu.

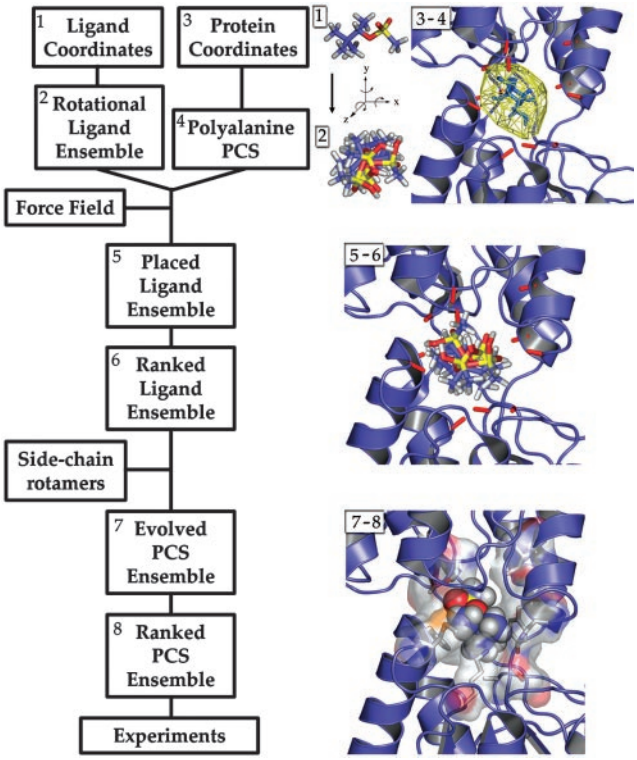

Fig. 2. Flow chart of the receptor design process. Numbers in the flow chart and molecular models correspond to steps described in the text: 1 and 2, rotational ligand ensemble; 3 and 4, truncated scaffold (red, alanine surface; yellow, convex hull); 5 and 6, placed ligand ensemble; 7 and 8, example of a complementary surface design.

IAF), *N*-((2-(iodoacetoxy)ethyl)-*N*-methyl)amino-7-nitrobenz-2-oxa-1,3-diazole (NBDE), and Acrylodan were purchased from Molecular Probes. JPW4039, JPW4042, and JPW4045 were gifts from Les Loew (University of Connecticut, Storrs).

Computational Design. The ReceptorDesign component of the DEZYMER program (11) was used to generate designs of mutant receptors. The design process consists of eight steps (Fig. 2). In step 1, the internal degrees of freedom within the ligand are sampled to identify low-energy ligand conformations (the internal ligand ensemble, ILE). A single, minimum-energy conformer of the PMPA *R*-isomer was used in this study. In step 2, a rotational ligand ensemble (RLE) is prepared in the absence of protein coordinates, sampling Eulerian rotations around the three principal molecular axes of the ligand (2.5° intervals, $\approx 10^6$ poses). In step 3, a pocket for the new binding site is identified by using the original ligand to locate the layer of residues that are in direct van der Waals or hydrogen bonding contact (the primary complementary surface, PCS). In step 4, residues in the PCS (excepting glycines or prolines) are replaced with alanine, generating a truncated protein scaffold representing a PCS for which no sequence has been determined. In step 5, the RLE is placed on each point of a cubic grid (0.5-Å spacing) within the convex hull, which envelops the ligand van der Waals surface. In step 6, a placed ligand ensemble (PLE) is constructed by selecting ranked members from these RLEs that are sterically compatible with the truncated scaffold, and confined within the convex hull (>90% of ligand atoms). In step 7, for each of top 10,000 docked ligands (selected from the PLE by choosing ligands with the fewest interactions with the truncated scaffold), a PCS is calculated. In this calculation, a side-chain rotamer library (an expanded version of ref. 37, containing 6,122 rotamers) representing all possible mutations (except cysteine or proline) and side-chain conformations is placed at all positions in the PCS, and a sequence corresponding to the global minimum energy of a pairwise-decomposed potential function is identified by a dead-end elimination algorithm (38). This potential function is based on a semiempirical force field that includes a modified Lennard–Jones potential to represent "fuzzy" van der Waals interactions (11, 21, 38) [parameters for amino acids and PMPA taken from CHARM22 (39) or a universal force field approximation (40, 41), respectively], an explicit geometry-dependent hydrogen-bonding term (11, 21, 38), a continuum solvation term to represent the hydrophobic effect with terms favoring or disfavoring burial of polar or nonpolar groups (11, 21, 38), and a linear term to account for differences in side-chain entropy ($E_s = wRT\ln N$, where N is the number of free torsions in the side-chain, and w is a weight; typically 1.0). Electrostatic contributions were not included in the calculations. The search algorithm maintains the ligand hydrogen bond inventory, selecting complementary sequences with minimal unsatisfied hydrogen bonds between ligand and protein. All PMPA oxygens were classified as hydrogen bond acceptors. In step 8, the predicted designs were ranked by four independent criteria: van der Waals contacts, hydrogen bonding energies between protein and ligand, the number of unsatisfied ligand hydrogen bonds, and exposed cavities within the binding pocket. Suitable designs were selected by taking the intersection of the top 10% of each ranked list. This linear optimization method optimizes fitness functions with components of different magnitudes and ranges. The final choice is based on visual inspection of the molecular models. The algorithm described here is similar to the one used previously (11) with enhanced ligand sampling (steps 5 and 6) and introduction of the final selection by linear optimization (step 8). The calculations were parallelized at steps 4 and 6, and carried out on a BEOWULF cluster of 20 1.7-GHz processors in ≈2 days per combination of scaffold and ligand.

Construction of Designed Receptor Proteins. Mutations were introduced into the RBP and GBP genes by using overlap extension PCR (42). A single cysteine was introduced in each of the constructs (RBP, Cys-265; GBP, Cys-112) for covalent attachment of a fluorescent reporter (19). Constructs were cloned with a C-terminal decahistidine tag in a pET21a expression vector by using 5′ *Xba*I and 3′ *Eco*RI restriction sites. Mutations were confirmed by DNA sequencing; expressed proteins were confirmed by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry. Proteins were purified by using immobilized metal affinity chromatography and labeled with reporter fluorophores (19). For GBP designs, all buffers contained 1 mM $CaCl_2$.

Ligand-Binding Analysis. Ligand binding was measured by direct titration into a solution of a covalently labeled protein (10–100 nM), monitoring changes in fluorescence emission intensity at 25°C (19).

Results

Computational Design. The binding pockets of RBP (PDB code 2dri) (17) and GBP (PDB code 2gbp) (43) were redesigned to bind PMPA by the ReceptorDesign component of DEZYMER (11), with 11 and 12 residues forming the primary complementary surface (PCS) in each receptor, respectively. The algorithm uses the three-dimensional structure of a protein to predict sequences and structures of binding sites that are complementary to a docked ligand (Fig. 2). A combinatorial search procedure simultaneously optimizes sequence choice and ligand docking to identify mutations that form complementary surfaces. Three RBP and 12 GBP designs were constructed by site-directed mutagenesis, and their ligand-binding properties were determined (Tables 1 and 2 and Fig. 3).

Each design corresponds to a separate PCS and a distinct orientation of the docked PMPA molecule. In all cases PMPA is sequestered within the binding site, with no direct contact with bulk solvent. In the majority of the designs, the methyl phosphonate

Table 1. Sequence and binding properties of the designed receptors derived from GBP

Design	Complementary surface sequence*												PMPA†‡ K_d, μM	PA† K_d, mM	MP† K_d, mM	Fluorophores§	ΔI_{std}¶ (PMPA)
	10_I	14_I	16_I	91_I	92_I	152_{II}	154_{II}	158_{II}	183_{II}	211_{II}	236_{II}	256_{II}					
*wt*GBP	Y	D	F	N	K	H	D	R	W	N	D	N	NB			IAF	0
PG4	**Q**	***K***	F	**A**	**S**	***N***	***N***	**I**	***H***	**S**	**A**	**S**	0.7	0.04	14	IAF	0.041 (−)
PG4_256F	**Q**	***K***	F	**A**	**S**	***N***	***N***	**I**	***H***	**S**	**A**	**F**	0.4	NB	12	IAF	0.015 (−)
PG5‖	**Q**	***K***	F	**S**	**S**	**N**	***S***	**S**	***Y***	**S**	**A**	**S**	NB				0
PG6‖	**K**	**H**	***H***	**A**	**S**	**N**	***S***	**I**	**F**	**S**	**A**	***S***	NB				0
PG7‖	**Q**	**S**	**L**	**V**	**S**	**N**	**S**	***Q***	**F**	***S***	**A**	**H**	NB				0
PG8‖	**Q**	***K***	F	**A**	**S**	**N**	***S***	**S**	***Y***	**S**	**A**	**S**	NB				0
PG9	**Q**	***K***	F	**A**	**S**	***N***	***H***	**I**	***H***	**S**	**A**	**S**	3.4	0.3	10	IAF	0.11 (+)
PG10	**K**	**M**	F	**A**	**S**	**N**	**Y**	**I**	***K***	***S***	**A**	***S***	0.44	1.5	12	IAF	0.22 (−)
PG11**	**K**	**S**	***H***	**V**	**S**	**N**	***S***	***Q***	**F**	**S**	**A**	**S**					
PG12	**Q**	***K***	F	**S**	**S**	***N***	***S***	**I**	***H***	**N**	**A**	**S**	0.25	0.8	9.5	IAF	0.15 (+)
PG12_256F	**Q**	***K***	F	**S**	**S**	***N***	***S***	**I**	***H***	**N**	**A**	**F**	0.11	1.3	11	IAF	0.23 (−)
PG14††	**M**	***R***	F	***S***	**S**	***N***	**H**	**K**	**F**	**S**	**A**	**H**					
PG17	**Q**	***K***	F	**S**	**S**	**N**	***S***	**I**	***Y***	**N**	**A**	**S**	5	NB	140	NBDE	0.059 (−)
PG18	**Q**	***K***	F	**A**	**S**	**N**	***S***	**I**	***Y***	**S**	**A**	**S**	10	8.7	NB	NBDE	0.014 (−)

*Amino acids are given in one-letter abbreviations. Subscripts indicate the location of the residue: I, N-terminal domain; II, C-terminal domain. Bold, mutation from wild type; italicized, hydrogen bond to PMPA.
†NB, no binding as determined by a fluorescence change at maximum ligand concentrations (PMPA, 10 mM; PA, 100 mM; MP, 1,000 mM). Blank entry, not determined.
‡Affinities were determined by using a racemic mixture of PMPA. The receptors were designed for the *R*-isomer.
§Fluorophores (IAF, NBDE, and Acrylodan) used to test the design for binding to PMPA; fluorophore used in determining affinity is presented.
¶Fluorescence emission intensity change in the presence of saturating ligand (ΔI_{std} as defined in ref. 19).
‖No ligand-mediated change in fluorescence upon addition of PMPA (10 mM).
**Protein precipitated.
††No expression.

group points out toward the solvent. In the case of the PG10 design in GBP, however, this group is oriented inwards (Fig. 4). In all designs, the hydrogen-bonding potential of both phosphonate anionic oxygens, as well as the phosphoester oxygen, are satisfied.

The majority of the designs were built in GBP (Table 1), and were selected from the top 50 ranked designs (Fig. 4), sampling both low- and high(er)-energy designs. The 12 PCS residues of the GBP designs can be divided into three groups according to the sequence diversity observed within the family of designs: constant (92_I, 152_{II}, 236_{II}), highly conserved (211_{II}, 256_{II}), and variable (10_I, 14_I, 16_I, 91_I, 154_{II}, 158_{II}, 183_{II}). The constant and highly conserved positions all differ from the wild-type protein. Two of the three constant residues arise from a change in function between the designs and the wild-type receptor. In wild-type GBP, Lys-92_I and His-152_{II} form hydrogen bonds to glucose. In most designed PMPA receptors, Ser-92_I and Asn-152_{II} do not interact with the ligand (in PG12 Asn-152_{II} forms an additional hydrogen bond with PMPA) but participate in a hydrogen-bonding network connecting the N- and C-terminal domains. This network may function as a "latch" that stabilizes the closed form (Fig. 3 *a* and *b*). The third constant residue (Ala-236_{II}) is constrained by steric differences between glucose and PMPA. In wild-type GBP, Asp-236_{II} forms a hydrogen bond to glucose; in all designs, the PMPA position precludes choice of any amino acid but alanine or glycine at this position. The highly conserved positions 211_{II} (Ser or Asn) and 256_{II} (Ser or His) also have switched from ligand binding (Asn-211_{II} and Asn-256_{II} interact with the O3 and O4 glucose hydroxyls, respectively) to structural functions. In 11 designs, Ser-211_{II} forms a hydrogen bond with the main-chain carbonyl of Val-235_{II}; in three designs, Asn-211_{II} interacts both with the amide proton of Met-214_{II} and the carbonyl of His-183_{II}. In the majority of the designs, Ser-256_{II}

Table 2. Sequence and binding properties of the designed receptors derived from RBP

Design	Complementary surface sequence*											PMPA†‡ K_d, μM	Fluorophore§	ΔI_{std}¶ (PMPA)
	13_I	15_I	16_I	89_I	90_I	137_H	141_{II}	164_{II}	190_{II}	215_{II}	235_{II}			
*wt*RBP	N	F	F	D	R	A	R	F	N	D	Q	NB		0
PR8	**S**	**A**	***S***	***S***	**D**	**A**	**M**	***K***	***K***	**A**	**S**	0.068	JPW4042	0.082 (+)
PR9‖	N	F	**L**	***N***	**D**	***S***	**M**	**H**	***S***	**A**	**S**	NB		0
PR11‖	N	F	**L**	***N***	**D**	***S***	**M**	**S**	***S***	**A**	**S**	NB		0
PR8_235A	**S**	**A**	***S***	***S***	**D**	A	**M**	**K**	***K***	**A**	**A**	0.069	JPW4039	0.08 (−)
PR8_235L	**S**	**A**	***S***	***S***	**D**	A	**M**	**K**	***K***	**A**	**L**	0.064	JPW4039	0.07 (−)
PR8_235F	**S**	**A**	***S***	***S***	**D**	A	**M**	**K**	***K***	**A**	**F**	0.045	JPW4039	0.063 (−)

*Amino acids are given in one-letter abbreviations. Subscripts indicate the location of the residue: I, N-terminal domain; II, C-terminal domain; H, hinge. Bold, mutation from wild type; italicized, hydrogen bond to PMPA.
†NB, no binding as determined by a fluorescence change at maximum ligand concentrations (PMPA, 10 mM; PA, 100 mM; MP, 1000 mM). Blank entry, not determined.
‡Affinities were determined by using a racemic mixture of PMPA. The receptors were designed for the *R*-isomer.
§Fluorophores (JPW4039, JPW4042, and JPW4045) used to test the design for binding to PMPA; fluorophore used in determining affinity is presented.
¶Fluorescence emission intensity change in the presence of saturating ligand (ΔI_{std} as defined in ref. 19).
‖No ligand-mediated change in fluorescence upon addition of PMPA (10 mM).

BIOCHEMISTRY

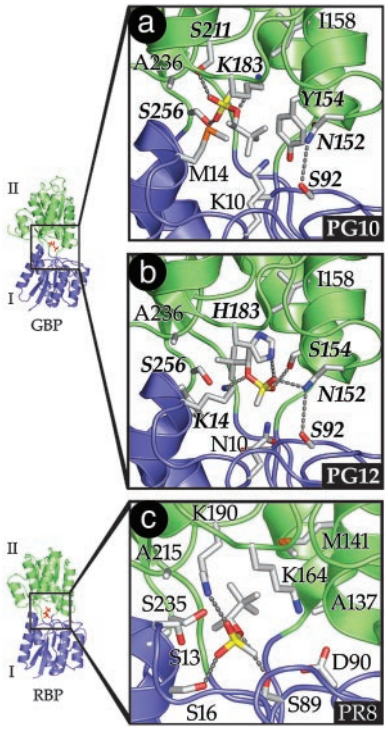

Fig. 3. Structures of GBP and RBP (domain I, blue; domain II, green) with computational models of representative designs (protons are not shown). (*a*) GBP design PG10. (*b*) GBP design PG12. (*c*) RBP design PR8. Residues selected for alanine-scanning mutagenesis are italicized.

forms a hydrogen bond with Gln-261_{II} outside the PCS (with the exception of PG10, where Ser-211_{II} forms a hydrogen bond to PMPA).

The designs leave a cavity between Ser-256_{II} and the PMPA pinacolyl group. The penalty for solvent accessibility of the hydrophobic ligand moiety apparently was insufficient to overcome the reward for forming the interresidue hydrogen bond. We constructed additional point mutations at position 256_{II} in designs PG4 and PG12 to fill this cavity (PG4_256F and PG12_256F; Table 1).

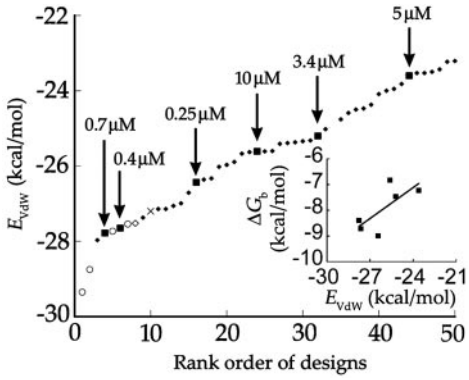

Fig. 4. Selection of GBP designs. ■, Ligand-mediated fluorescent response (experimentally observed affinities are indicated); ◆, not tested; ○, no fluorescent response; x, no protein expression; ◇, protein precipitation. Designs were chosen from a final list of candidates by using a linear optimization procedure that selected a subset corresponding to the intersection of the top 20% ligand van der Waals energy, 50% ligand H-bond energy, with all H-bonds satisfied and with solvent-accessible surface areas <15 Å^2 (see *Materials and Methods*). The designs are shown ranked by the van der Waals energy (E_{vdw}) of the interaction between ligand and receptor, which is a measure of close packing. (*Inset*) Correlation between the experimentally determined PMPA affinities and E_{vdw} for the tested designs.

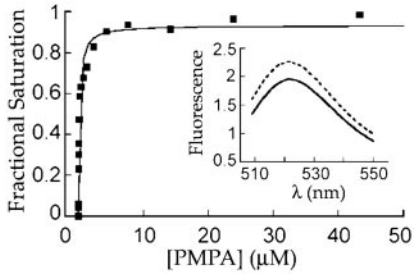

Fig. 5. Fluorescent response of IAF-labeled PG12 upon titration with PMPA. (*Inset*) Emission spectra of protein in the absence (solid line) and presence of 0.5 mM PMPA (dashed line).

Sequences at the variable positions are diverse: on average, 33% of the residues differ among the designs, reflecting alternative ways for providing hydrogen bonds and hydrophobic surfaces. The designs vary in their PCS positions at which hydrogen-bonding side-chains are placed.

The three designs constructed in RBP also exhibit variations in sequence diversity and residue function switching. In PR8, Ser-235_{II} is associated with a defect analogous to Ser-256_{II} in GBP. Ser-235_{II} makes no direct contacts with PMPA, but forms a hydrogen bond with the hydroxyl of Ser-103_{II}, resulting in a cavity near the pinacolyl group. To fill this cavity, additional point mutations were constructed in the RBP design PR8 at position 235_{II} (Table 2).

Construction of Designed Receptors. All three RBP designs and 10 of the 12 primary GBP designs expressed soluble protein; one GBP design did not express, whereas another precipitated upon purification. Several of the mutants are less thermostable than the parent proteins (GBP, 58°C; RBP, 60°C), having thermostabilities that range between 32°C and 58°C as determined by thermal denaturation, monitoring circular dichroism (21).

Ligand-Binding Properties. Of the 18 fluorescent conjugates prepared by labeling with thiol-reactive fluorophores at Cys-256 (RBP) or Cys-112 (GBP), 12 show changes in fluorescence upon addition of PMPA (Fig. 5). Neither wild-type RBP nor GBP conjugates respond to PMPA.

Observed PMPA affinities range from 68 nM (PR8) to 10 μM (PG18) (Tables 1 and 2). Some of the cavity-filling mutations constructed at position 235_{II} in RBP show improvements in affinity. Phenylalanine at position 235_{II} increases the affinity of the receptor for PMPA (K_d = 45 nM), whereas Ala-235_{II} or Ile-235_{II} have no effect (Table 2). The equivalent mutation at position 256_{II} in GBP (PG4_256F, K_d = 0.4 μM; PG12_256F, K_d = 0.11 μM) has similar effects on binding (Table 1).

The ligand-binding specificity of two designs was tested by measuring affinities for IMPA (Fig. 1), the hydrolysis product of the nerve agent sarin (14). PG10 and PG12 bind IMPA ≈10-fold less tightly that PMPA (K_d = 7 μM and 2 μM, respectively), indicating significant discrimination between the aliphatic groups of the two molecules.

The affinities of the designs for pinacolyl alcohol (PA) and methyl phosphonate (MP), representing the aliphatic and hydrophilic moieties of PMPA, respectively (Fig. 1), were determined (Table 1). The K_d values of the receptors for PA and MP are 10^2- to 10^4-fold and 10^4- to 10^5-fold higher than those for PMPA, respectively. A coupling energy (44), ΔG_c, can be defined as: $\Delta G_c = \Delta G_{b,PMPA} - (\Delta G_{b,PA} + \Delta G_{b,MP})$, where $\Delta G_{b,PMPA}$, $\Delta G_{b,PA}$, and $\Delta G_{b,MP}$ are the binding energies ($RT \ln K_d$) for PMPA, PA, and MP, respectively. Favorable interfragment interactions result in $\Delta G_c < 0$, unfavorable $\Delta G_c > 0$. Analysis of fragment binding is typically

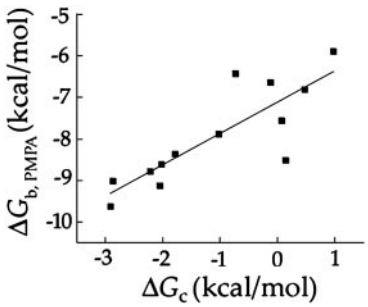

Fig. 6. Correlation between experimentally determined fragment coupling energy, ΔG_c, and the affinity for PMPA, $\Delta G_{b,PMPA}$.

used to assess strain or entropic factors within a ligand (44). Here, ΔG_c values between designs are interpreted as strain within the designed proteins, reflecting differences in the structural complementarity between a design and its bound ligand. Fig. 6 reveals a positive correlation between ΔG_c and the affinity for PMPA: as ΔG_c decreases, $\Delta G_{b,PMPA}$ becomes more favorable. Decreases in fragment strain therefore correlate with increased receptor affinities, and indicate differences in the complementarity of the designed surfaces.

Alanine-Scanning Mutagenesis. The contributions of specific interactions were tested by alanine scanning mutagenesis in two designs, PG10 and PG12 (Table 3) (45), which bind PMPA in opposite orientations (Fig. 3 *a* and *b*). In the PG10 design (MP moiety points inwards) mutation of predicted hydrogen bonds to an anionic oxygen (O1, PG10_S211A) or the phosphoester oxygen (O3, PG10_K183A) results in a 2.1 and 2.4 kcal/mol loss of binding energy, respectively, consistent with typical hydrogen bonding contributions (46). Loss of the predicted interaction between Ser-256_{II} and the other anionic oxygen has no appreciable effect (O2, PG10_S256A), potentially indicating that this hydrogen bond is absent. Ser-256_{II} is also predicted to form a hydrogen bond with Gln-261_{II}. The two interactions therefore may compete rather than coexist. In the PG12 design (MP moiety points outwards), loss of predicted hydrogen bonds to the anionic oxygens (O1, PG12_N152A; O2, PG12_S154A; O1, PG12_H183A) results in a 2–3 kcal/mol loss of binding energy, consistent with the model (Table 3).

Van der Waals interactions were also investigated. In PG10, Tyr-154_{II} interacts with the pinacolyl moiety of PMPA and hydrogen bonds to Thr-110_{II}. Loss of these predicted interactions decreases binding by 2.4 kcal/mol (Table 3). Furthermore, binding of PA, but not MP, is affected, consistent with the orientation of PMPA in the model. Similarly, in PG12, Asn-211_{II} forms van der Waals interactions with the pinacolyl moiety and hydrogen bonds to the backbone carbonyl of position 214_{II}. Loss of these predicted interactions (PG12_N211A) results in a decreased affinity for PMPA, but to a lesser extent (0.9 kcal/mol) than is observed for the Tyr-154_{II} in PG10. Again, as expected, PA, but not MP, binding is affected.

Alanine-scanning mutagenesis has also demonstrated that the interdomain latch, contributed by constant residues Ser-92_I and Asn-152_{II}, is important for binding (Table 3). Mutations of either residue decrease binding, as expected for the removal of an interaction that stabilizes the closed state (47, 48).

The Ser-256_{II}Ala mutation in PG12 exhibits the largest change in affinity (4 kcal/mol) (Table 3). This residue is not predicted to interact directly with PMPA, instead it hydrogen bonds to Gln-261_{II}, leaving a cavity. Enlargement of this putative cavity in the alanine mutation is predicted to trap water near the hydrophobic pinacolyl moiety, thereby decreasing the affinity for PMPA. Loss of PA and retention of MP binding in this mutant is observed and consistent with this interpretation.

Discussion

We report the computational design of a drastic switch in the ligand-binding specificity of RBP and GBP to bind PMPA, an organophosphate surrogate of the nerve agent soman, instead of their cognate sugars. Although PMPA is similar in size to glucose and ribose, it is structurally distinct from either sugar. We have

Table 3. Alanine point mutants in PG10 and PG12

	PMPA		PA		MP		
Design*	K_d, μM	$\Delta\Delta G^{\dagger}$, kcal/mol	K_d, mM[‡]	$\Delta\Delta G^{\dagger}$, kcal/mol	K_d, mM[‡]	$\Delta\Delta G^{\dagger}$, kcal/mol	Interaction[§]
PG10	**0.45**		**1.3**		**12**		
PG10_S92A	19	2.2	0.29	−0.87	NB	NB	Latch
PG10_N152A	56	2.9	0.92	−0.19	12	0	Latch
PG10_Y154A	23	2.4	8.7	1.1	9.1	−0.17	van der Waals contact to pinacolyl group, and hydrogen bond to T110
PG10_K183A	12	2.0	0.32	−0.81	17	0.21	Hydrogen bond to O3
PG10_S211A	16	2.1	1.8	0.21	11	−0.08	Hydrogen bond to O1
PG10_S256A	0.6	0.02	1.6	0.12	11	−0.08	Hydrogen bond to O2 and Q162
PG12	**0.3**		**0.8**		**9.5**		
PG12_S92A	2	1	1	0.2	11	0.1	Latch
PG12_N152A	7	2	NB		8.5	0.06	Hydrogen bond to O1, and latch
PG12_S154A	90	3	NB		9.8	0.02	Hydrogen bond to O2
PG12_H183A	30	3	NB		9.1	−0.02	Hydrogen bond to O1
PG12_N211A	0.9	0.8	2	0.5	9.0	−0.03	Hydrogen bond to M214 carbonyl, van der Waals contact to pinacolyl group
PG12_S256A	200	4	NB		17	0.4	Hydrogen bond to Q261

*Labeled with IAF.

[†]$\Delta\Delta G = -RT \ln K_d^{Design}/K_d^{Mutant}$.

[‡]No binding (NB) at maximal ligand concentrations (PA, 100 mM; MP, 1,000 mM).

[§]See Fig. 1 for oxygen numbering scheme.

previously shown that RBP can be redesigned to bind metals (21), TNT, or L-lactate, and GBP can be redesigned to bind L-lactate (11). Together, these experiments demonstrate the ability to change the ligand-binding specificity of proteins for chemically disparate ligands.

Of the computer-generated designs, ≈50% show PMPA-mediated changes in fluorescence of the covalently coupled reporter groups (57% if designs that do not express or that precipitate are discounted). This success rate represents a lower bound, because false negatives can arise if the equilibrium between the open and closed states is sufficiently altered to preclude their interconversion, or if the fluorophore no longer interacts differentially with these two conformations.

PMPA affinities of the designed receptors range from 45 nM to 10 μM. RBP and GBP bind their cognate sugars with 0.2 μM and 0.5 μM affinities, respectively (19). Empirical limits have been established for the ligand affinities of naturally evolved proteins (49). For PMPA, this limit ranges from ≈2 nM to ≈1 μM. The affinities of many designs reported here fall within this range and rival or exceed those of the parent receptors.

Selected designs sample both high- and lower-ranked candidates. Designs selected from the top 20 exhibit higher affinities for PMPA than those selected from lower-ranked designs (Fig. 4). Analysis of the affinities for PA and MP suggests that the designed receptors differ in the strain they impose on the ligand (Fig. 6) (44).

The effects of individual alanine mutations on PMPA binding in designs PG10 and PG12 are mostly consistent with the predicted interactions. Furthermore, the designed receptors distinguish steric differences between the aliphatic moieties of PMPA and IMPA (Fig. 1). We therefore conclude that predicted molecular models of the designs are largely correct.

The designs contain defects, indicating that the computational design methods require further improvements. Virtually all designs have a cavity between the protein and bound ligand in the vicinity of the hinge region. This cavity defect is likely to be a consequence of inaccurate modeling of relative contributions by hydrogen bonds, polar group burial, solvent accessibility, and omission of electrostatic contributions. Nevertheless, the experimentally validated ligand-binding properties of the designs reported here demonstrate that even relatively simple representations of atomic interactions are sufficiently powerful to capture dominant effects of biomolecular recognition in design calculations.

The designed PCS has fewer residues that make direct contacts with the ligand than those in the wild-type receptors. Consequently, a significant fraction of the side-chains switch function from ligand binding in the wild-type receptor to a structural role in the designed receptors and lack sequence diversity. However, the residues that interact directly with the ligand are highly diverse and depend on the orientation of the bound ligand. Thus even in this small set of designs, significant diversity in structure and sequence is observed, suggesting that solutions to the design problem are highly degenerate. These observations presumably reflect a fundamental characteristic of protein sequences, because potential diversity is an essential prerequisite for the genesis of function by the random processes of organic evolution (50).

The receptors described here can function as reagentless fluorescent biosensors for PMPA with a lower detection limit of ≈4 nM (≈1 ppb). Given the structural similarities between soman and PMPA, the designed receptors are likely to bind soman with affinities similar to those of PMPA. The detection limit is probably sufficient for the development of stand-off or postincident detectors of soman (51), and rivals the lower limits of current methods (32, 36). Unlike acetylcholinesterase-based assays, the designed receptors described here do not rely on the presence of soman, which rapidly degrades to form PMPA (14). Other techniques require several components and longer preparation, incubation, and detection times (31–36, 52). A reagentless fluorescence biosensor has significant advantages, such as rapidity of the fluorescent response, reversibility, and simplicity. However, the molecular recognition element in a deployable biosensor must be sufficiently robust to withstand field conditions. The designed receptors reported here do not yet meet this standard, because their thermostability is still not sufficiently high. Nevertheless, computationally designed receptors represent a promising step in the development of a class of biosensors for the rapid, continuous, and accurate detection of nerve agents.

We thank L. Loew for the styryl dyes, G. Shirman for assistance with mutagenesis and protein chemistry, and W. L. DeLano for providing the PYMOL software. This work was supported by a grant from the Defense Advanced Research Project Agency.

1. Arnold, F. H. (2001) *Nature* **409,** 253–257.
2. Lazar, G. A., Marshall, S. A., Plecs, J. J., Mayo, S. L. & Desjarlais, J. R. (2003) *Curr. Opin. Struct. Biol.* **13,** 513–518.
3. Barker, P. D. (2003) *Curr. Opin. Struct. Biol.* **13,** 490–499.
4. Bolon, D. N., Voigt, C. A. & Mayo, S. L. (2002) *Curr. Opin. Chem. Biol.* **6,** 125–129.
5. Kraemer-Pecore, C. M., Wollacott, A. M. & Desjarlais, J. R. (2001) *Curr. Opin. Chem. Biol.* **5,** 690–695.
6. Gordon, D. B., Marshall, S. A. & Mayo, S. L. (1999) *Curr. Opin. Struct. Biol.* **9,** 509–513.
7. Mendes, J., Guerois, R. & Serrano, L. (2002) *Curr. Opin. Struct. Biol.* **12,** 441–446.
8. Desjarlais, J. R. & Clarke, N. D. (1998) *Curr. Opin. Struct. Biol.* **8,** 471–475.
9. Kuhlman, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L. & Baker, D. (2003) *Science* **302,** 1364–1368.
10. Havranek, J. J. & Harbury, P. B. (2003) *Nat. Struct. Biol.* **10,** 45–52.
11. Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. (2003) *Nature* **423,** 185–190.
12. Kortemme, T., Ramirez-Alvarado, M. & Serrano, L. (1998) *Science* **281,** 253–256.
13. Dahiyat, B. I. & Mayo, S. L. (1997) *Science* **278,** 82–87.
14. Munro, N. B., Talmage, S. S., Griffin, G. D., Waters, L. C., Watson, A. P., King, J. F. & Hauschild, V. (1999) *Environ. Health. Perspect.* **107,** 933–974.
15. Tam, R. & Saier, M. H., Jr. (1993) *Microbiol. Rev.* **57,** 320–346.
16. Mowbray, S. L. & Cole, L. B. (1992) *J. Mol. Biol.* **225,** 155–175.
17. Bjorkman, A. J., Binnie, R. A., Zhang, H., Cole, L. B., Hermodson, M. A. & Mowbray, S. L. (1994) *J. Biol. Chem.* **269,** 30206–30211.
18. Quiocho, F. A. & Ledvina, P. S. (1996) *Mol. Microbiol.* **20,** 17–25.
19. de Lorimier, R. M., Smith, J. J., Dwyer, M. A., Looger, L. L., Sali, K. M., Paavola, C. D., Rizk, S. S., Sadigov, S., Conrad, D. W., Loew, L., *et al.* (2002) *Protein Sci.* **11,** 2655–2675.
20. Benson, D. E., Conrad, D. W., de Lorimier, R. M., Trammell, S. A. & Hellinga, H. W. (2001) *Science* **293,** 1641–1644.
21. Dwyer, M. A., Looger, L. L. & Hellinga, H. W. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 11255–11260.
22. Benson, D. E., Haddy, A. E. & Hellinga, H. W. (2002) *Biochemistry* **41,** 3262–3269.
23. Marvin, J. S. & Hellinga, H. W. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 4955–4960.
24. Hellinga, H. W. & Marvin, J. S. (1998) *Trends Biotechnol.* **16,** 183–189.
25. National Research Council Committee on Alternative Chemical Demilitarization Technologies, Longwell, J. P., National Research Council Board on Army Science and Technology & National Research Council Commission on Engineering and Technical Systems (1993) *Alternative Technologies for the Destruction of Chemical Agents and Munitions* (Natl. Acad. Press, Washington, DC).
26. Kortepeter, M. G., Cieslak, T. J. & Eitzen, E. M. (2001) *J. Environ. Health* **63,** 21–24.
27. Hay, A. (2000) *Med. Confl. Surviv.* **16,** 37–41.
28. Noort, D., Hulst, A. G., Platenburg, D. H., Polhuijs, M. & Benschop, H. P. (1998) *Arch. Toxicol.* **72,** 671–675.
29. Vermillion, W. D. & Crenshaw, M. D. (1997) *J. Chromatogr. A* **770,** 253–260.
30. Kukkonen, J., Riikonen, K., Nikmo, J., Jappinen, A. & Nieminen, K. (2001) *J. Hazard. Mater.* **85,** 165–179.
31. Lee, W. E., Thompson, H. G., Hall, J. G. & Bader, D. E. (2000) *Biosens. Bioelectron.* **14,** 795–804.
32. Erhard, M. H., Jungling, A., Schoneberg, T., Szinicz, L. & Losch, U. (1993) *Arch. Toxicol.* **67,** 220–223.
33. Nassar, A. E., Lucas, S. V., Jones, W. R. & Hoffland, L. D. (1998) *Anal. Chem.* **70,** 1085–1091.
34. Wang, J., Chatrathi, M. P., Mulchandani, A. & Chen, W. (2001) *Anal. Chem.* **73,** 1804–1808.
35. Wang, J., Pumera, M., Collins, G. E. & Mulchandani, A. (2002) *Anal. Chem.* **74,** 6121–6125.
36. Driskell, W. J., Shih, M., Needham, L. L. & Barr, D. B. (2002) *J. Anal. Toxicol.* **26,** 6–10.
37. Lovell, S. C., Word, J. M., Richardson, J. S. & Richardson, D. C. (2000) *Proteins* **40,** 389–408.
38. Looger, L. L. & Hellinga, H. W. (2001) *J. Mol. Biol.* **307,** 429–445.
39. MacKerell, A. D., Jr., Bashford, D., Bellott, M., Dunbrack, R. L., Jr., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., *et al.* (1998) *J. Phys. Chem. B* **102,** 3586–3616.
40. Mayo, S. L., Olafson, B. D. & Goddard, W. A., III (1990) *J. Phys. Chem.* **94,** 8894–8909.
41. Rappe, A., Casewit, C. J., Colwell, K. S., Goddard, W. A., III, & Skiff, W. M. (1992) *J. Am. Chem. Soc.* **114,** 10024–10035.
42. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. (1989) *Gene* **77,** 51–59.
43. Vyas, N. K., Vyas, M. N. & Quiocho, F. A. (1988) *Science* **242,** 1290–1295.
44. Jencks, W. P. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 4046–4050.
45. DeLano, W. L. (2002) *Curr. Opin. Struct. Biol.* **12,** 14–20.
46. Fersht, A. (1999) *Structure and Mechanism in Protein Science* (Freeman, New York).
47. Marvin, J. S. & Hellinga, H. W. (2001) *Nat. Struct. Biol.* **8,** 795–798.
48. Millet, O., Hudson, R. P. & Kay, L. E. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 12700–12705.
49. Kuntz, I. D., Chen, K., Sharp, K. A. & Kollman, P. A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9997–10002.
50. White, S. H. (1994) *Annu. Rev. Biophys. Biomol. Struct.* **23,** 407–439.
51. Institute of Medicine Committee on R & D Needs for Improving Civilian Medical Response to Chemical and Biological Terrorism Incidents & National Research Council Board on Environmental Studies and Toxicology (1999) *Chemical and Biological Terrorism: Research and Development to Improve Civilian Medical Response* (Natl. Acad. Press, Washington, DC).
52. Jenkins, A. L., Uy, O. M. & Murray, G. M. (1999) *Anal. Chem.* **71,** 373–378.