Biophysical Journal Volume 91 July 2006 401–410

A Computational Protocol for the Integration of the Monotopic Protein Prostaglandin H2 Synthase into a Phospholipid Bilayer

Philip W. Fowler and Peter V. Coveney
Centre for Computational Science, Department of Chemistry, University College London, London, United Kingdom

ABSTRACT Prostaglandin H2 synthase (PGHS) synthesizes PGH_2, a prostaglandin precursor, from arachidonic acid and was the first monotopic enzyme to have its structure experimentally determined. Both isozymes of PGHS are inhibited by nonsteroidal antiinflammatory drugs, an important class of drugs that are the primary means of relieving pain and inflammation. Selectively inhibiting the second isozyme, PGHS-2, minimizes the gastrointestinal side-effects. This had been achieved by the new PGHS-2 selective NSAIDs (i.e., COX-2 inhibitors) but it has been recently suggested that they suffer from additional side-effects. The design of these drugs only made use of static structures from x-ray crystallographic experiments. Investigating the dynamics of both PGHS-1 and PGHS-2 using classical molecular dynamics is expected to generate new insight into the differences in behavior between the isozymes, and therefore may allow improved PGHS-2 selective inhibitors to be designed. We describe a molecular dynamics protocol that integrates PGHS monomers into phospholipid bilayers, thereby producing in silico atomistic models of the PGHS system. Our protocol exploits the vacuum created beneath the protein when several lipids are removed from the top leaflet of the bilayer. The protein integrates into the bilayer during the first 5 ns in a repeatable process. The integrated PGHS monomer is stable and forms multiple hydrogen bonds between the phosphate groups of the lipids and conserved basic residues (Arg, Lys) on the protein. These interactions stabilize the system and are similar to interactions observed for transmembrane proteins.

INTRODUCTION

Monotopic proteins are integral membrane proteins but, unlike transmembrane proteins, their polypeptide chains do not cross the phospholipid bilayer (1). To date only four monotopic proteins have had their structures experimentally determined; prostaglandin H2 synthase (PGHS) (2), squalene-hopene cyclase (3), monoamine oxidase (4), and fatty acid amide hydrolase (5). There is tremendous academic and industrial interest in monotopic proteins; for example, all four enzymes for which structures exist are important pharmaceutical drug targets.

PGHS (EC No. 1.14.99.1) catalyses the conversion of arachidonic acid, a 20-carbon fatty acid, to prostaglandin H_2 (PGH_2), the precursor of the prostaglandin class of local hormones. The reaction proceeds within PGHS in two steps at spatially distinct active sites: arachidonate and two molecules of oxygen are reacted together at the cyclooxygenase (COX) site to form prostaglandin G_2 (PGG_2), which is then reduced to PGH_2 at the peroxidase site (6). PGHS is often called COX due to its cyclooxygenase function. We will not discuss the structure, biochemistry, and inhibition of PGHS in detail as there are many excellent reviews (7–11).

PGHS is interesting because not only is it a monotopic protein but also two genes encode similar PGHS enzymes (11). Additional splice variants of the PGHS-1 isozyme have been proposed, potentially increasing this number (12). PGHS-1 is constitutively expressed and synthesizes prostaglandins involved in homeostasis, for example maintaining the mucosal lining of the stomach (13). PGHS-2 is induced and rat models have demonstrated that this enzyme is involved in local pain and inflammation responses (14,15). The PGHS enzymes have also been implicated in other human pathologies, for example in various cancers (10).

Nonsteroidal antiinflammatory drugs (NSAIDs), such as aspirin, ibuprofen, and flurbiprofen, bind within the cyclooxygenase active site, limiting the production of PGH_2 and therefore inhibiting the action of the enzyme. The ulcerogenic and renal side-effects of NSAIDs are caused by inhibition of PGHS-1 and the majority of classical NSAIDs (e.g., aspirin and ibuprofen) are either not selective or inhibit PGHS-1. Consequently there has been a tremendous effort in the last decade to design an NSAID that is specific to PGHS-2 and therefore has reduced side-effects, while maintaining the desired analgesic and antiinflammatory actions. This culminated in 1999 in the introduction of several so-called COX-2 inhibitors, most notably celecoxib (Celebrex, Pfizer) and rofecoxib (Vioxx, Merck). Interest in these important clinical inhibitors has increased recently due to the withdrawal of rofecoxib in 2004 by its manufacturer because of a higher incidence of myocardial infarction during an extended clinical trial. The mechanism of this side-effect is not known. A recent population study by Hippisley-Cox and Coupland (16) suggested that other NSAIDs, including ibuprofen, may also suffer from the same side-effect, albeit to differing degrees.

Designing an isozyme-specific inhibitor is an extremely difficult task because the isozymes are structurally very similar.

Submitted November 15, 2005, and accepted for publication January 26, 2006.

Address reprint requests to Professor Peter V. Coveney, Tel.: 44-20-7679-4850; E-mail: p.v.coveney@ucl.ac.uk.

0006-3495/06/07/401/10 $2.00

doi: 10.1529/biophysj.105.077784

For example, the sequence similarity between PGHS-1 and PGHS-2 is 60–65% within the same species (10) and the root mean-square deviation (RMSD) over the C_α atoms is 0.9–1.0 Å when comparing the different sheep PGHS-1 and mouse PGHS-2 x-ray crystal structures. The PGHS active site, as defined by those residues in close contact with the bound substrate, is highly conserved with only a single amino acid difference (I523V) between isozymes. We will follow the convention of using the amino acid numbering of sheep PGHS-1. The less bulky side chain of valine compared to isoleucine permits ligands to access an additional side-pocket seen in the x-ray crystallographic structures of PGHS-2 (17). The COX-2 inhibitors exploit this pocket and mutagenesis experiments have demonstrated that this mutation is important in producing the selectivity of these new inhibitors (18–20).

The x-ray crystallographic structure of a monotopic protein is a static snapshot of the protein at cryogenic temperatures and contains little dynamical information. Kinetic studies have revealed that the dynamics of PGHS is important for its inhibition by NSAIDs (21,22). Studying the dynamics of PGHS-1 and PGHS-2 is likely to lead to additional insight beyond that gained when comparing static x-ray crystallography structures and therefore could inform future drug design. This study is the first step toward analyzing such dynamical differences between PGHS-1 and -2.

We use large-scale classical molecular dynamics (MD) to study the dynamics of this system. MD is constrained by the short timescales, typically tens of nanoseconds, that it can access relative to those of more general biological interest, but it allows insight to be gained into the dynamics (and therefore behavior) of these proteins that is not possible by experiment. Hypotheses may then be generated which can be tested experimentally in an iterative process.

It is not known what aspects of the system will be important in determining the dynamical differences (if any) between the PGHS isozymes and it is therefore prudent to include a phospholipid bilayer, since this forms many interactions with PGHS. Unfortunately, it is difficult to build atomistic models of monotopic proteins anchored to a membrane as, unlike transmembrane proteins, there are no clear transmembrane units (e.g., α-helices or a β-barrel) with which to position the protein relative to the membrane. The purpose of this article is to outline a protocol for integrating a PGHS monomer into a phospholipid bilayer. We expect a PGHS dimer to integrate in the same way as the monomer. We shall present evidence to demonstrate that the protein is correctly inserted before making some concluding remarks.

METHOD

In this section we describe the structure of PGHS, define our protocol for integrating a PGHS monomer with a phospholipid bilayer, and provide details about the MD algorithm used.

Structure of prostaglandin H2 synthase

PGHS is situated on both the luminal side of the endoplasmic reticulum and on the nuclear membrane (23,24). X-ray crystallography indicates that the protein is a dimer, is mainly α-helical, and comprises 550–553 residues per monomer. Each monomer has three domains (see Fig. 1): an epidermal-growth factor (EGF)-like domain (residues 33–72), the catalytic domain (residues 117–586), and the membrane-binding domain (MBD, residues 73–116). A PGHS monomer is not biologically active, and therefore, if we drew any conclusions from the simulations of the monomer about the dynamical differences between the two isozymes, we would first have to first establish that the structure and dynamics of an integrated PGHS monomer are similar to that of an integrated PGHS dimer.

The MBD is composed of four α-helices, *A*, *B*, *C*, and *D*, the first three roughly forming three sides of a square with *D* connecting this motif to the catalytic domain. Picot et al. (2) hypothesized that these three short helices lie in the plane of the membrane near one interface, form numerous interactions with the top leaflet of the bilayer, and thereby bind the protein to the membrane. This was supported by surface plots of the hydrophobicity

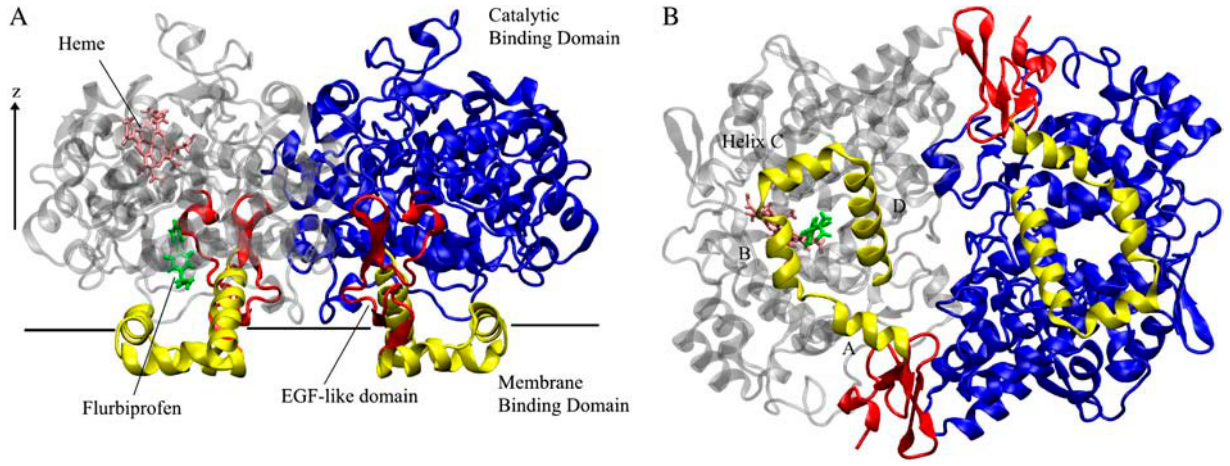

FIGURE 1 The secondary and tertiary structure of *ovine sp*. prostaglandin H2 synthase-1 dimer. A side view of PGHS is drawn in panel *A* and the different domains are labeled. The EGF-like domains are colored red, the membrane binding domain yellow, and the catalytic domains, blue and gray. Flurbiprofen and the heme cofactor are colored green and pink, respectively, in the left-hand (*gray*) monomer. All views are with respect to the postulated plane of the membrane, which is shown as a dark line. The view from underneath PGHS is drawn in panel *B* and the four different helices (*A–D*) of the MBD are labeled.

of the protein and comparison with other globular heme-containing peroxidases lacking the EGF and MBD domains, notably mammalian myeloperoxidase (10). The hypothesis was subsequently confirmed by a variety of photolabeling, mutagenesis (23,25), and fluorescent protein fusion experiments (24).

The hydrophobic substrate, arachidonic acid, is partitioned into the membrane and, it is assumed, enters the enzyme directly via the MBD. After cyclooxygenation, the product, PGG_2, is expelled from the COX site of PGHS and then binds to a peroxidase-active site where the final conversion to PGH_2 takes place. Differences between PGHS-1 and PGHS-2 that may be exploited in the design of inhibitors could potentially be in any one of these steps. We assume here that it is most likely that differences will occur in the interactions between the drug and the enzyme; hence, we study flurbiprofen bound to the COX active site of PGHS-1 and -2. However, it is possible that other important differences may exist—for example, in the mechanism of entry of the substrate into each isozyme. This has been studied by Molnar et al. (26) using steered molecular dynamics.

Integration protocol

We will now describe the protocol used to integrate a monomer PGHS with a phospholipid bilayer. It was decided to integrate monomers rather than dimers to limit the size the system; however, we expect this approach to apply to PGHS dimers also. The protocol exploits the force exerted by the vacuum created when a number of lipids are removed from the bilayer directly beneath PGHS to integrate the enzyme rapidly into the membrane. This approach was first used by Nina et al. (27) to integrate the MBD into a small patch of lipids and we extend it here to an entire PGHS monomer. This has the advantage that we do not need any restraints on the protein and, by virtue of increased computer processor speeds and improved algorithms, we were able to evolve our models for 15 ns compared to the 1 ns performed by Nina et al. (27).

The membrane plug-in to VMD1.8.3 (28) was used to generate a patch of bilayer comprising 214–234 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPC) lipids, POPC being the most common constituent of the endoplasmic reticulum (29). The protein was oriented such that helices *A*, *B*, and *C* of the MBD were in the plane of the phospholipid bilayer. Defining the *z* axis as perpendicular to the phospholipid bilayer with its origin at the center of the bilayer (see Fig. 1), the number of lipids within 2 Å of the protein was calculated as a function of *z*. The center of mass of the protein was then moved to a *z* coordinate (48–50 Å) such that it was in close contact with 17 lipids.

Deciding how many lipids to remove is a packing problem; in principle, one should remove the number of lipids (11) equivalent to the cross-sectional area of the MBD to ensure the membrane is minimally perturbed. This, however, does not allow the enzyme to be positioned very close to the bilayer as the MBD does not fit well into the cavity produced due to discrete size effects. An additional six lipids were removed to allow the MBD to better fit the cavity and we expect the perturbation introduced by the integration of PGHS into the membrane to dominate any potential perturbation in the curvature of the membrane due to the removal of these additional lipids. Any effect is further reduced by using a large patch of POPC lipids.

Having removed these 17 POPC lipids to create the cavity beneath the protein, the protein was solvated and then neutralized by adding counterions; care was taken to ensure no water molecules entered this cavity. The potential energy of the system was then minimized and the water and side chains relaxed before the system was thermalized up to physiological temperature (310 K) over 0.5 ns. During the warming, harmonic restraints in the *z* direction were applied to selected heavy atoms within the headgroups of the phospholipids. The magnitude of the restraints was decreased and removed before a Berendsen barostat was applied to maintain the system at 1 atm pressure. A Langevin thermostat was used to maintain the temperature at 310 K and production runs were 15 ns long. This duration is typical for current transmembrane protein simulations (30).

TABLE 1 PGHS systems simulated

Identifier	Species	Complex	PDB structure used	Size (atoms)
aPGHS-1	Sheep	Apo	1prh (2)	103,683
aPGHS-2	Mouse	Apo	5cox (39)	101,815
cPGHS-1	Sheep	Flurbiprofen	1eqh (40)	111,022
cPGHS-2	Mouse	Flurbiprofen	3pgh (39)	107,353

Systems studied

Table 1 lists the four PGHS monomers studied and their respective Protein Data Bank (PDB) codes (31). The heme group was included but the proteins were not glycosylated, as it has been shown that this is not necessary for their function (10).

The flurbiprofen-bound PGHS-2 system (referred to as cPGHS-2) is typical and contains a PGHS monomer, 1 flurbiprofen molecule, 1 heme group, 6 sodium cations, 215 POPC molecules, and 23,193 TIP3P water molecules, with typical dimensions ~80 Å × 80 Å × 150 Å.

Algorithms

We used NAMD2.5, a parallel classical molecular dynamics algorithm to minimize, warm, equilibrate, and evolve the dynamics of our system (32). NAMD has excellent scaling capabilities and therefore allows rapid computation using either regular high performance computing (HPC) or grid computing resources. For example, using the IA-64 Linux cluster at NCSA (part of the US TeraGrid—www.teragrid.org), we are able to simulate 1 ns in 6.25 wallclock hours when running on 192 processors. The compute nodes of the UK National Grid Service (www.ngs.ac.uk) and HPCx (www.hpcx.ac.uk) yielded comparable performance.

The standard protein and lipid CHARMM27 force field was used, which includes parameters for POPC and heme (33). The unknown bonding parameters and all the partial charges for flurbiprofen were determined using an MP2/SBK ab initio approach. SHAKE was used to admit a 2 fs integrator timestep (34) and the particle-mesh Ewald method was used to compute the electrostatic forces (35). Conformations were saved every 2–5 ps and van der Waals interactions were cut off at 12 Å with a switching distance of 10 Å. Steered molecular dynamics simulations used a spring constant of 0.486 nN/Å and a bead velocity of 5×10^{-5} Å per timestep. VMD1.8.3 (28) was used to perform all analysis and to produce the illustrations within this article.

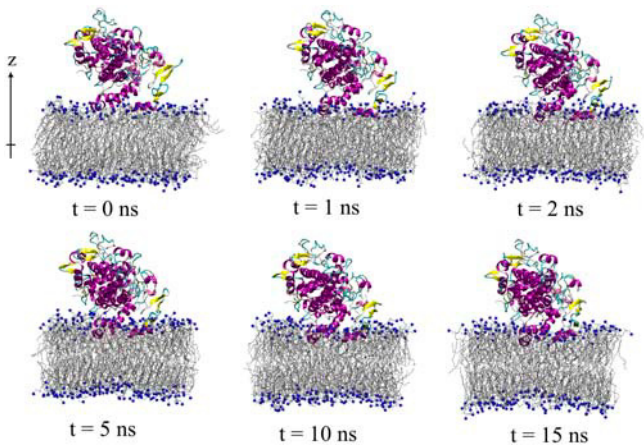

FIGURE 2 Snapshots of the monomeric PGHS-1 system at different times. The nitrogen atoms of the choline groups of the POPC lipids are drawn as blue spheres to allow easy identification of the lipid-water interface.

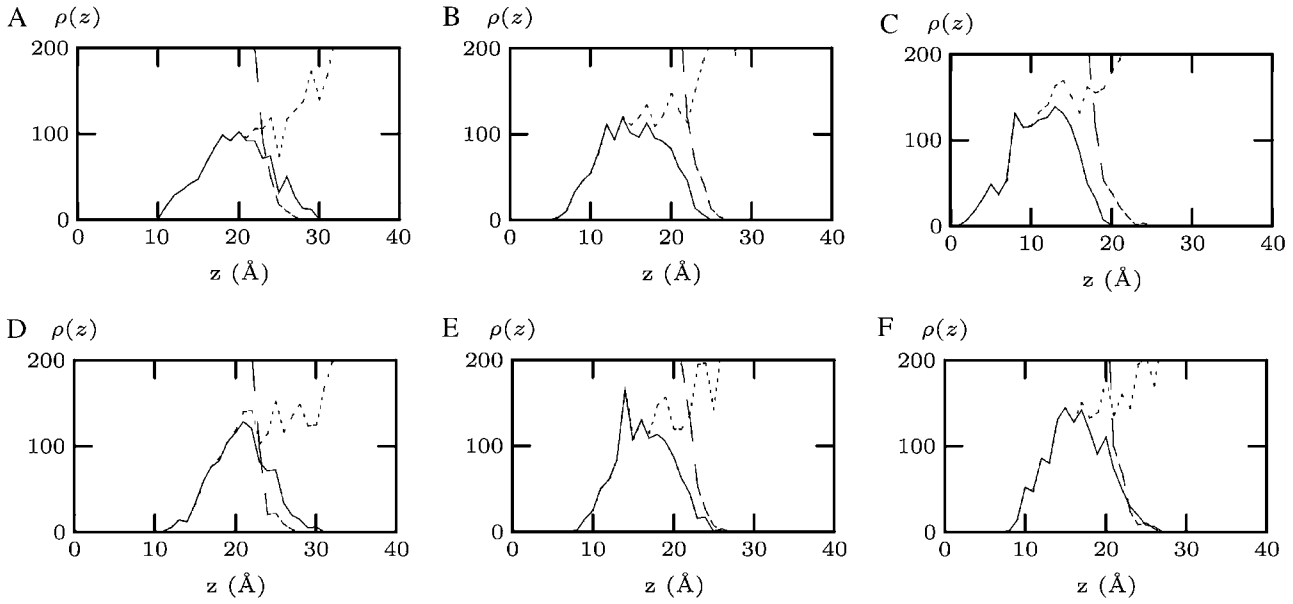

FIGURE 3 The atomic densities, $\rho(z)$, as a function of z for the membrane binding domain (*solid line*), the whole enzyme (*dotted line*), and the phospholipid bilayer (*dashed line*) for both PGHS-1 and PGHS-2 show the integration of the protein into the bilayer. Panels *A–C* and *D* and *E* show the progressive integration of PGHS-1 and PGHS-2, respectively, at $t = 0$, $t = 5$ ns, and $t = 14.75$ ns. The value $\rho(z)$ is averaged over a 250 ps window and is measured in kg m^{-3}.

RESULTS

The aim of the integration protocol is to integrate PGHS into a phospholipid bilayer to ensure that the environment, and therefore the dynamics, of the protein is more representative of that in vivo. Given the assumptions inherent in our computational models and the relatively short timescales that we can access by MD, we cannot prove definitively that the protein has correctly integrated in the bilayer, but we can present evidence indicating it is very likely to be correctly integrated.

In this section we will describe results that demonstrate that:

1. The protein becomes well integrated into the bilayer.
2. The protocol is repeatable.
3. PGHS is stable throughout.
4. It is bound to the membrane.
5. PGHS interacts with individual phospholipids in a manner reminiscent of the interactions formed by transmembrane proteins.

Integration

Snapshots of cPGHS-1 at different times (see Fig. 2) show PGHS integrating with the bilayer. A more quantitative measure of integration is shown in Fig. 3, *A–C*; here, the atomic densities, $\rho(z)$, of the membrane-binding domain (MBD), the whole enzyme, and the lipid bilayer are averaged over 250 ps windows and plotted as a function of z, where z is the distance from the center of the phospholipid bilayer. This shows that the MBD (and the whole protein) sinks ~5 Å into the membrane in the first 5 ns with little change in the following 10 ns.

Following the analysis of Deol et al. (36) on transmembrane proteins, we define the number of close contacts, $c(t)$, between the protein and the lipid bilayer as the number of lipid atoms within 3.5 Å of the protein. This is plotted in Fig. 4 and the number of lipid atoms in contact with cPGHS-1 increases from ~50–400. As expected, if we separate the interactions made between the protein and headgroups and tails of the lipids, the majority of this increase is due to additional interactions with lipid tails as the protein sinks into the bilayer and the lipids rearrange (data not shown). If we assume that a protein is integrated when a plateau in $c(t)$ is attained, then this metric and our previous analyses indicate that cPGHS-1 is integrated (and therefore equilibrated) after 5 ns.

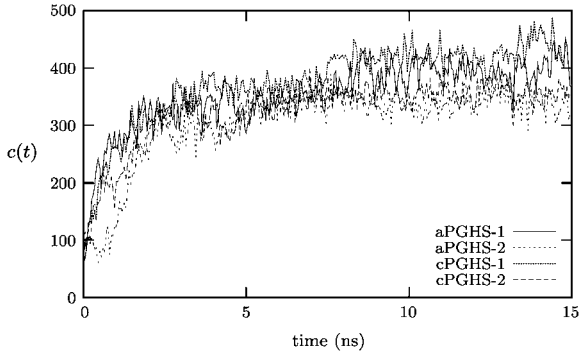

FIGURE 4 The number of close contacts as a function of time, $c(t)$, for each of the four systems studied (see Table 1), reaches a plateau after ~5 ns.

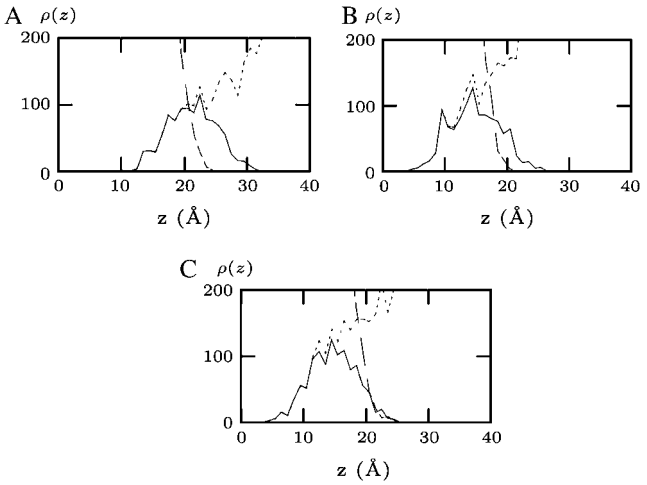

FIGURE 5 Atomic densities, $\rho(z)$, as a function of z for the MBD (*solid line*), the whole enzyme (*dotted line*), and the phospholipid bilayer (*dashed line*) for the restrained integration of aPGHS-1. Densities are plotted at different times: $t = 0$ (panel *A*), $t = 5$ ns (panel *B*), and $t = 14.75$ ns (panel *C*). The value $\rho(z)$ is averaged over a 250 ps window and is measured in kg m^{-3}.

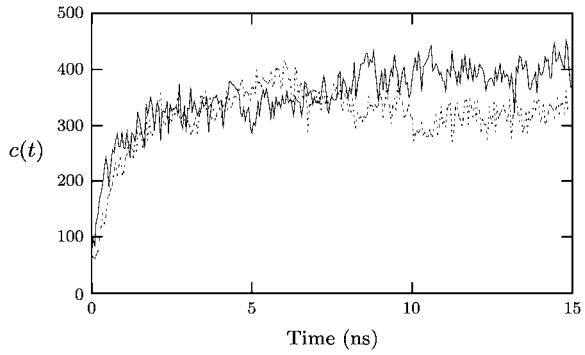

FIGURE 6 The number of close contacts as a function of time, $c(t)$, for the restrained integration of apo PGHS-1. The restrained and unrestrained simulation are drawn with a solid line and a broken line, respectively.

Repeatability

We expect PGHS to integrate into the membrane independently of the initial conditions. If we assume that the force exerted due to the cavity dominates any differences between the two isozymes or between the apo and flurbiprofen-bound forms, then we also expect different isozymes of PGHS to integrate in the same way. The repeatability of the protocol is tested in two ways: 1), the integration of the four different PGHS systems given in Table 1 is compared; and 2), a modified protocol, with alternating restraints, is compared to the unmodified protocol for aPGHS-1.

Examining $c(t)$ for all four systems indicates that each forms similar numbers of close contacts with lipids and that each isozyme is integrated after ~5 ns (Fig. 4). The distributions of atomic densities, $\rho(z)$ for cPGHS-1 and cPGHS-2 are compared in Fig. 3 and are also similar, although $\rho(z)$ suggests that the MBD of cPGHS-2 does not appear as well integrated after 15 ns as cPGHS-1. This is not supported by examining $c(t)$.

To further check the sensitivity of the protocol, an additional integration simulation was run with the backbone of aPGHS-1 and the headgroups of the phospholipids alternatively restrained for 0.1–0.5 ns by an harmonic potential. The distributions of the densities for this restrained simulation are shown in Fig. 5; as expected, the MBD is less well integrated after 5 ns; however, after 15 ns, the MBD is entirely within the bilayer. The number of close contacts, $c(t)$, is plotted in Fig. 6 and is similar for both the restrained and unrestrained simulations. The differences in $c(t)$ between the restrained and unrestrained simulations are within the variation seen for different systems in Fig. 4. The protocol is therefore repeatable; it is moderately insensitive to changes and different PGHS systems are integrated to the same degree.

Stability

The rapid integration of PGHS into the phospholipid bilayer might be expected to significantly perturb the conformation of the protein. Again, we assume that the differences in sequence and structure between the four systems are dominated by the force exerted by the cavity. Fig. 7 plots the C_α RMSD for the four systems against their respective x-ray crystallographic structures (see Table 1). After an initial increase, the RMSD for each system fluctuates with a maximum of 3.0 Å for aPGHS-1 and a minimum of 1.8 Å for

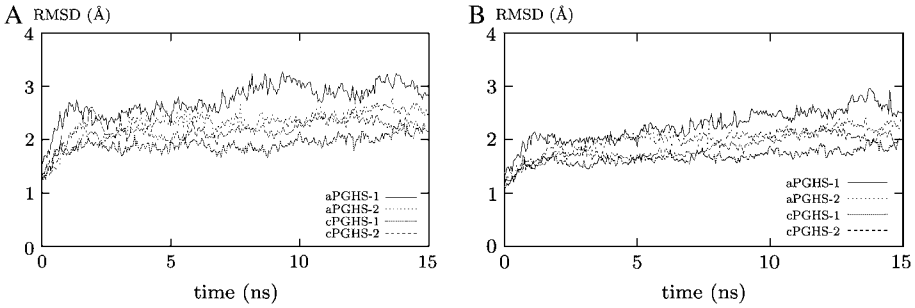

FIGURE 7 The root mean-square deviations as a function of time for each of the four systems studied. The differences between the RMSDs for the different systems shown in panel *A* are reduced by fitting on each domain separately (*B*), thereby removing any interdomain motion. Either RMSD measure for all four systems is <3 Å throughout and therefore their structures remain similar to their x-ray crystallographic structures.

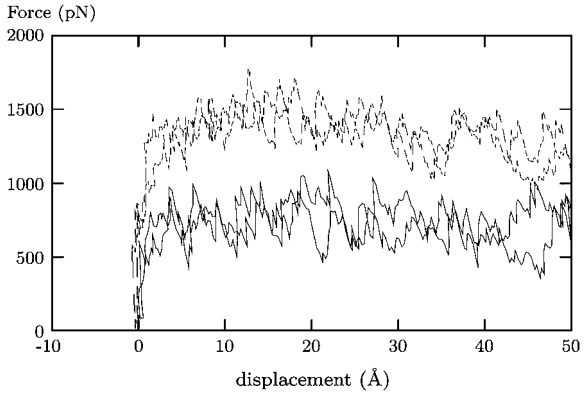

FIGURE 8 The force applied to PGHS for the bound (*dashed lines*) systems and the unbound (*solid lines*) systems, showing how a significantly larger force is required to move the bound system the same distance as the unbound system. Due to the fluctuations in the force, each bound and unbound case was repeated twice.

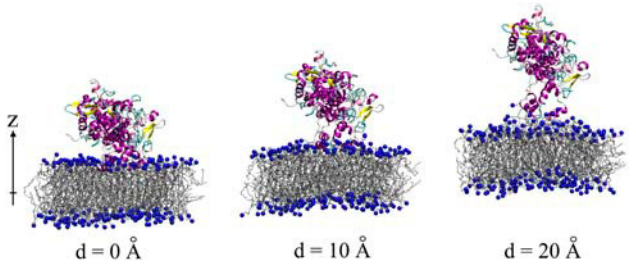

FIGURE 9 Snapshots from a steered molecular dynamics simulation as PGHS is pulled free of the phospholipid bilayer at three different values of the displacement, *d*. Note the deformation of the lipid bilayer.

cPGHS-1. This range of RMSD values for a protein of this size indicates that the protein is stable throughout the simulations and therefore also during the integration protocol. To assess how much of the difference in behavior between systems was due to interdomain motion, the mean-square deviation of each domain was calculated and then combined to give an RMSD of the whole protein, which ignores differences in domain orientation (RMSD′). This is plotted in Fig. 7 *B* and all four systems have more similar values of RMSD′ (1.5–2.5 Å) than RMSD, indicating that some of the structural changes are due to interdomain motions.

To verify that the dynamics of a PGHS monomer is not likely to be significantly different to the dimer, the average structures were compared to the x-ray crystallographic structures and the RMSD for each residue was calculated. Examining the RMSD of the surface residues of the proteins indicates that the dimer interface, with the exception of a short α-helix (residues 363–370), has similar RMSD values to the remainder of the surface residues (data not shown).

It is usual to assume that a protein has equilibrated when a plateau in RMSD is attained. This implies that all four systems have equilibrated after 2.5 ns. This is half the time estimated for the proteins to integrate with the phospholipid bilayer and is consistent with a rapid sinking of the protein into the bilayer followed by a rearrangement of the lipids around the MBD.

Binding

To test whether the protein is bound to the phospholipid bilayer after the integration process (as defined by the first 5 ns of each production trajectory), we applied a force to the aPGHS-1 system to try to unbind it. We used the steered molecular dynamics functionality of NAMD (37). A bead is connected by a spring to the center of mass of the protein, then pulled with a constant velocity in the *z* direction and the force in the spring recorded. To control for the viscous forces experienced by the protein due to the water, we repeated the experiment for a simple solvated aPGHS-1 system. Finally, both systems were rerun from a different initial conformation.

Fig. 8 shows that the force required to move the membrane-bound aPGHS-1 protein a fixed distance was significantly higher than the force required to move the solvated aPGHS-1 protein. This implies that the aPGHS-1 protein has formed interactions with the bilayer preventing its free movement. Examining the trajectories shows that the membrane-bound protein deforms under the tension and the MBD remains attached to the bilayer for some time, partially deforming the bilayer itself before being pulled free (Fig. 9). In the next section we will study the interactions the MBD forms with the phospholipid bilayer.

```
PGHS1-sheep          EIWTWLRTTLRPSPSFIHFMLTHGRWLWDFVN-ATFIRDTLMRLVLTVRSNLIP
PGHS2-mouse          EFLTRIRLLLRPTPNTVHYILTHFKGVWNIVNNIPFLRSLIMKYVLTSRSYLID
PGHS1-human          GLWTWLRNSLRPSPSFTHFLLTHGRWFWEFVN-ATFIREMLMRLVLTVRSNLIP
PGHS1-rabbit         DLWTWLRSSLRPSPTFVHYLLTHVRWFWEFVN-ATFIRDTLMRLVLTVRSNLIP
PGHS1-mouse          EIWTWLRNSLRPSPSFTHFLLTHGYWLWEFVN-ATFIREVLMRLVLTVRSNLIP
PGHS1-rat            EIWTWLRSSLRPSPSFTHFLLTHGYWIWEFVN-ATFIREVLMRLVITVRSNLIP
PGHS1-rainbow_fish   EFWTRVHQQLKPSPDVVHYILTHFHWLWNLIN-RTFMRDWLMRVVLTVRSNLIP
PGHS2-rat            EFLTRIRLLLRPTPNTVHYILTHFKGVWNIVNNIPFLRNSIMKYVLTSRSHLID
PGHS2-guinea_pig     EFLTRIRLLLRPTPNTVHYILTHFKGVWNIVNNIPFLRNAIMIYVLTSRSHLID
PGHS2-rabbit         EFLTRIRLLLRPTPDTVHYILTHFKGVWNIVNSIPFLRNSIMKYVLTSRSHMID
PGHS2-horse          EFLTRIRLFLRPTPNTVHYILTHFKGVWNIVNSFPFLRNAVMKYVLVSRSHLIE
PGHS2-american_mink  EFLTRVRLLLRPTPNTVHYILTHFKGVWNIVNRIPFLADVIMKYVRTSRSHCIE
PGHS2-cow            EFLTRIRLLLRPTPNTVHYILTHFKGVWNIVNRISFLRNMIMKYVLTSRSHLIE
PGHS2-sheep          EFLTRIRLLLRPTPDTVHYILTHFKGVWNIVNRISFLRNMIMKYVLTSRSHLIE
PGHS2-pig            EFLTRIRLFLRPTPNTVHYILTHFKGVWNIVNNIPFLRNAIMKYVLISRSHLID
PGHS2-human          EFLTRIRLFLRPTPNTVHYILTHFKGFWNVVNNIPFLRNAIMSYVLTSRSHLID
PGHS2-chicken        EFFTWLRLILRPTPNTVHYILTHFKGVWNIINNISFLRDTIMRYVLTSRSHLID
PGHS2-rainbow_fish   EFLTWIRISLRPAPNTIHYILTHYKGLWNVINRITFVRNAIMSYVLTPRSHLVD
```

FIGURE 10 Multiple-alignment of PGHS-1 and PGHS-2 primary sequences (MBD region only). Note that the first two sequences are those with x-ray crystallographic structures and are those studied in this article. Basic residues (*W*,*Y*) and aromatic amphipathic residues (*R*,*K*) are colored purple and yellow, respectively.

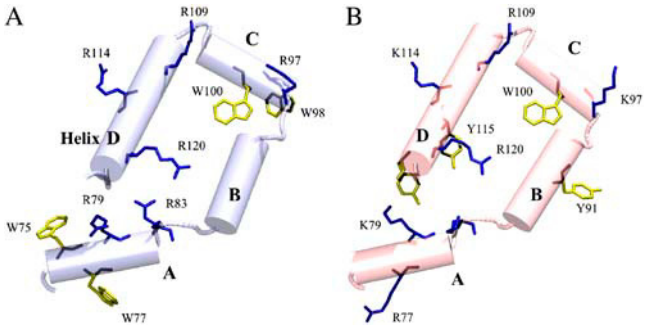

FIGURE 11 The distribution of basic and aromatic amphipathic residues on the membrane binding domain (MBD). oPGHS-1 is on the left, mPGHS-2 is on the right. Basic residues (Arg, Lys) are shown in blue and aromatic amphipathic residues (Trp, Tyr) are shown in yellow. Aromatic and amphipathic residues have been shown to interact with individual phospholipids and stabilize transmembrane proteins in membranes.

Protein-lipid interactions

Although the sequence conservation between PGHS-1 and -2 is 60–65%, it is only 38% within the membrane binding domain (MBD) (23). Despite this there are many conserved or semiconserved residues (see Fig. 10), notably amphipathic aromatic (Tyr, Trp) and basic residues (Arg, Lys). For a more complete sequence alignment, see the review by Simmons et al. (11). These residues are found on the helices as well as within the interior of the MBD and on the surface (Fig. 11). The high degree of conservation in a region of overall low homology implies that these residues have a specific function.

Conserved bands of aromatic amphipathic and basic residues are observed for many different transmembrane proteins (38); these residues form stabilizing interactions with the phospholipid bilayer. For example, so-called snorkeling interactions between basic residues and the phosphate groups have been observed during computational studies of KcsA and OmpA (36). In our simulations, both basic and aromatic residues were observed forming close interactions with phospholipids after the integration of the protein into the phospholipid bilayer. We will concentrate on the hydrogen bonds formed by the basic residues since these are more easily quantified.

One example of an interaction between a basic residue and one or two phospholipids from each of the cPGHS-1 and cPGHS-2 simulations is illustrated in Figs. 12 and 13. Arginine is potentially very promiscuous as its guanidinium group allows it to form several hydrogen bonds (or a single ionic interaction) with several oxygens of one or more phosphate groups. For this example, the hydrogen bonds formed with Arg-114 persist for longer (up to 10 ns) than the hydrogen bond formed with Lys-114 for cPGHS-2. Finally, as we would expect, the hydrogen bonds occasionally swap from one oxygen atom on the phosphate group to another.

It is difficult to draw general conclusions from individual interactions, so we seek more global measures now. The total number of hydrogen bonds, $h(t)$, formed between the protein and the lipids is plotted in Fig. 14. With the exception of cPGHS-2, the number of hydrogen bonds formed increases in the first 5 ns to 5–10. For both isozymes, the basic residues make up the majority of the hydrogen bonds formed between the MBD and the phospholipid bilayer; however, there are differences due to the lack of arginines on the MBD of PGHS-1 (Figs. 10 and 11).

It has been suggested in the literature that the PGHS isozymes have significantly different free energies of binding to membranes (23). As we have seen, the guanidinium group of arginine allows it to form multiple simultaneous hydrogen bonds and therefore the substitution of arginines by lysines

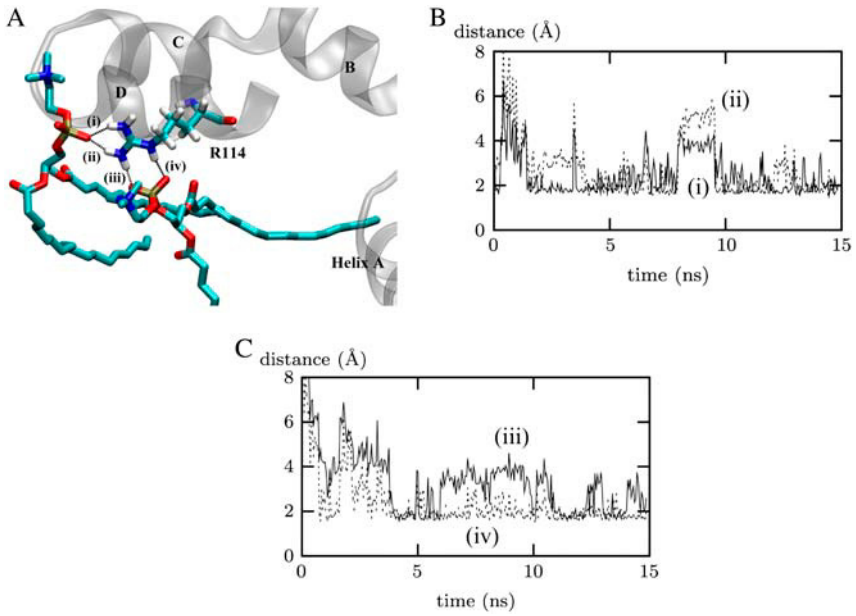

FIGURE 12 An example of a typical interaction between an arginine of cPGHS-1 and several phospholipids. Panel *A* shows Arg-114 of cPGHS-1 forming four hydrogen bonds, labeled *i–iv*, with two separate POPC lipids. The hydrogen-bonding distances for these bonds are plotted in panels *B* and *C*. The bonds persist for up to 10 ns and the guanidinium group bonds to different oxygens on the phospholipids. Hydrogen bonds are indicated by a dashed line and each is labeled. The lipids are drawn without hydrogens for clarity and only the MBD is drawn and the different helices are labeled.

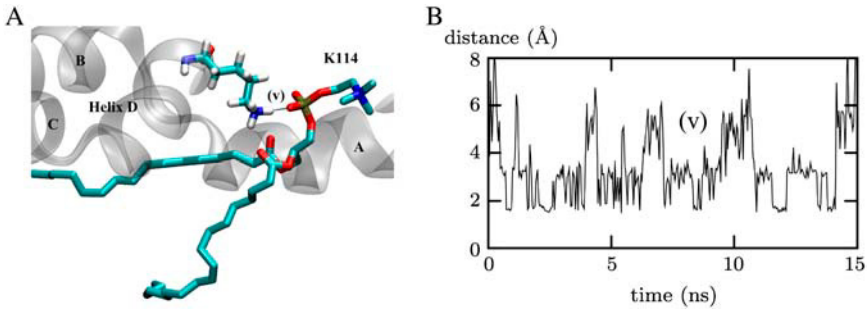

FIGURE 13 An example of a typical interaction between a lysine of cPGHS-2 and a phospholipid. Panel *A* shows Lys-114 of cPGHS-1 forming a single hydrogen bond, labeled (*v*), with the phosphate group of a POPC lipid. This hydrogen bond is intermittent and persists for no longer than 2 ns (panel *B*). The lipids are drawn without hydrogens for clarity and only the MBD is drawn and the different helices are labeled.

on the MBD of PGHS-2 could explain the difference between the binding free energies.

In all x-ray crystallographic structures, helix *B* is observed to lie slightly above helices *A* and *C* relative to the plane of the membrane (Fig. 1). In a previous study of the MBD, Nina et al. (27) calculated the free energy of solvation of the helices *A–D* using a mean field approach. Their results suggested that helix *B* should lie in the same plane as helices *A* and *C* and its reported position was therefore taken to be an artifact of the extraction and crystallization process. Helix *B*, however, remains above helices *A* and *C* in all our simulations (data not shown).

We suggest that this is because, unlike helices *A* and *C*, helix *B* has no basic residues that can form stabilizing snorkeling interactions. As a result, helix B integrates less deeply into the phospholipid bilayer than helix *A* or *C*. An atomistic effect of this nature would not be detected by a mean field approach; however, there is a histidine residue on helix *B* that we chose to make neutral in our simulations. Were this residue to be protonated, then one might expect helix *B* to form hydrogen bonds with a lipid phosphate group, become more integrated into the bilayer and therefore be drawn toward the plane of helices *A* and *C*. This could be tested by further simulations. We note that there is some limited evidence from mutagenesis experiments to support the hypothesis that helix *B* binds less strongly to the phospholipid bilayer than its neighbors, although this is not conclusive (23).

CONCLUSIONS

Monotopic proteins are an interesting and important class of proteins that merit investigation. In common with other integral membrane proteins, determining their structure and studying their function is difficult. Computational approaches, such as classical molecular dynamics, allow the scientist to gain insight into the behavior of monotopic proteins and from these studies suggest hypotheses that may be tested by experiment.

Although some aspects of how PGHS binds to the membrane are known, the precise orientation of the protein and the depth to which it sinks in the membrane is unknown. We have described a protocol to rapidly integrate a PGHS monomer into a phospholipid bilayer. The force exerted by the cavity produced when several lipids are removed from directly beneath the monomer is exploited to pull the protein into the membrane. This protocol is an extension of the earlier work carried out by Nina et al. (27). It would be interesting to examine whether this protocol could be applied to the integration of squalene-hopene cyclase, as it appears to bind to membranes in a similar manner to PGHS (3).

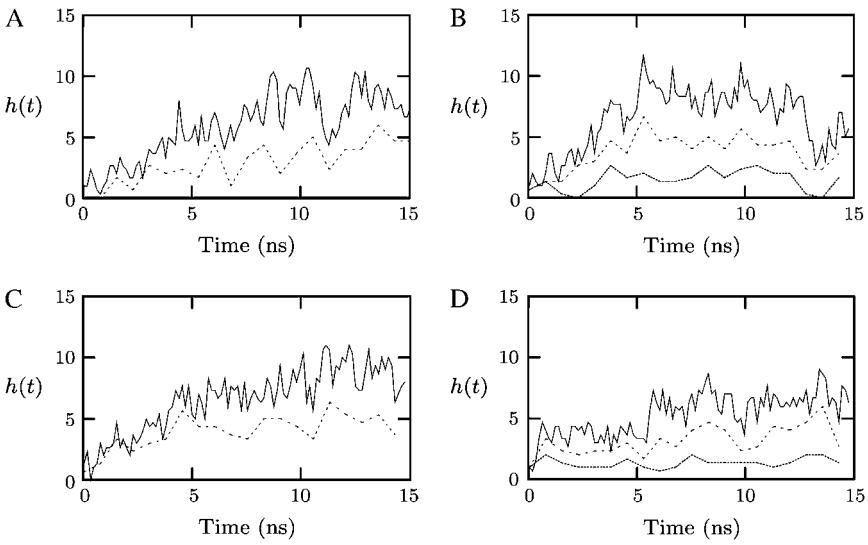

FIGURE 14 The total number, $h(t)$, of hydrogen bonds between the MBD and the lipid bilayer (*solid line*) for the four PGHS systems studied (see Table 1). The cumulative number of hydrogen bonds involving arginine (*dashed line*) and lysine (*dotted line*) are also plotted. A hydrogen bond is assumed to have formed if the distance between donor and acceptor is <3.5 Å and the deviation is $<30°$. The total number of hydrogen bonds for apo-PGHS-1 and -2 is plotted in panels *A* and *B*, and for $h(t)$ for flurbiprofen-bound PGHS-1 and -2 is plotted in panels *C* and *D*.

Our protocol was applied to both apo and flurbiprofen-bound PGHS-1 and PGHS-2 systems. For all four systems, the PGHS monomer sinks by ~5 Å into the bilayer and the number of close contacts made between the monomer and lipids increases significantly during the first 5 ns. The protocol therefore integrates the PGHS monomer with the phospholipid bilayer within the first 5 ns and is repeatable. The monomer is also stable throughout all our simulations. To demonstrate that, once integrated, the protein is bound, we used steered molecular dynamics to apply a force perpendicular to the plane of the membrane.

The most compelling evidence that the monomer is located at the correct depth in the bilayer is the formation of stabilizing snorkeling interactions between the conserved basic residues of the membrane binding domain and the phosphate groups of the lipids. Analogous interactions have been observed in simulations of transmembrane proteins (36). A series of mutagenesis experiments on the conserved basic residues of the membrane binding domain would help test our hypothesis that their function is to form stabilizing interactions. None of the mutagenesis experiments performed on the membrane-binding domain to date has investigated the role of the basic residues (23).

Depending on the motion, the relaxation times of phospholipids in bilayers can be much larger than the timescales we have studied here. This makes it more difficult to assess whether the entire system is equilibrated, particularly whether the membrane has adapted to the presence of the monotopic protein. We assume, however, that following the rapid integration of the PGHS monomers into the bilayer in the first 5 ns, the environment surrounding the enzyme more closely mimics that in vivo. The integration of PGHS with a phospholipid bilayer is the first step in assessing the dynamical differences between PGHS-1 and PGHS-2. In future publications we will report on the differences that we observe for both monomers and dimers, the hypotheses that are subsequently generated, and possible experimental tests of our theoretical findings.

We are grateful to Mark Sansom and his group for helpful discussions and to David Benoit for computing the CHARMM force-field parameters for flurbiprofen. HPC*x* generously allowed us to test as early adopters the Phase2 machine. Extensive use was also made of the computer nodes on the UK National Grid Service.

We thank the Engineering and Physical Science Research Council (EPSRC) (UK) for funding RealityGrid under grant No. GR/R67699, which provided access to the United Kingdom National Supercomputing facilities at HPC*x* and the Biotechnology and Biological Sciences Research Council (UK) for funding Integrative Biological Simulation under grant No. BBS/B/16011. HPC*x* also generously allowed us to test as early adopters the Phase2 machine. Access to the NCSA compute node of the US TeraGrid was provided through National Science Foundation (USA) NRAC and PACI grants MCA04N014 and ASC030006P, respectively. Extensive use was also made of the compute nodes of the UK National Grid Service. The PhD studentship of P.W.F. is funded by an EPSRC studentship associated with RealityGrid.

REFERENCES

1. Blobel, G. 1980. Intracellular protein topogenesis. *Proc. Natl. Acad. Sci. USA.* 77:1496–1500.
2. Picot, D., P. Loll, and R. Garavito. 1994. The x-ray crystal structure of the membrane protein prostaglandin H2 synthase-1. *Nature.* 367:243–249.
3. Wendt, K. U., K. Poralla, and G. E. Schulz. 1997. Structure and function of a squalene cyclase. *Science.* 277:1811–1815.
4. Binda, C., P. Newton-Vinson, F. Hubálek, D. E. Edmondson, and A. Mattevi. 2002. Structure of human monoamine oxidase-b, a drug target for the treatment of neurological disorders. *Nat. Struct. Biol.* 9:22–26.
5. Bracey, M. H., M. A. Hanson, K. R. Masuda, R. C. Stevens, and B. F. Cravatt. 2002. Structural adaptations in a membrane enzyme that terminates endocannabinoid signaling. *Science.* 298:1793–1796.
6. Hamberg, M., and B. Samulsson. 1967. On the mechanism of the biosynthesis of prostaglandins E_1 and $F_{1\alpha}$. *J. Biol. Chem.* 242: 5336–5343.
7. Smith, W. L., D. L. DeWitt, and R. M. Garavito. 2000. Cyclooxygenases: structural, cellular and molecular biology. *Annu. Rev. Biochem.* 69:145–182.
8. Kurumbail, R. G., J. R. Kiefer, and L. J. Marnett. 2001. Cyclooxygenase enzymes: catalysis and inhibition. *Curr. Opin. Struct. Biol.* 11:752–760.
9. Kulmacz, R. J., W. A. van der Donk, and A.-L. Tsai. 2003. Comparison of the properties of prostaglandin H synthase-1 and -2. *Progr. Lip. Res.* 42:377–404.
10. Garavito, R., and A. M. Mulichak. 2003. The structure of mammalian cyclooxygenases. *Annu. Rev. Biophys. Biomol. Struct.* 32:183–206.
11. Simmons, D. L., R. M. Botting, and T. Hla. 2004. Cyclooxygenase isozymes: the biology of prostaglandin synthesis and inhibition. *Pharmacol. Rev.* 56:387–437.
12. Warner, T. D., and J. A. Mitchell. 2002. Cyclooxygenase-3 (COX-3): filling in the gaps toward a COX continuum? *Proc. Natl. Acad. Sci. USA.* 99:13371–13373.
13. Goetzl, E. J., S. An, and W. L. Smith. 1995. Specificity of expression and effects of eicosanoids mediators in normal physiology and human disease. *FASEB J.* 9:1051–1058.
14. Masferrer, J. L., B. S. Zweifel, P. T. Manning, S. D. Hauser, K. M. Leahy, W. G. Smith, P. C. Isakson, and K. Seibert. 1994. Selective inhibition of inducible cyclooxygenase 2 *in vivo* is antiinflammatory and nonulcerogenic. *Proc. Natl. Acad. Sci. USA.* 91:3228–3232.
15. Seibert, K., Y. Zhang, K. M. Leahy, S. D. Hauser, J. L. Masferrer, W. Perkins, L. Lee, and P. C. Isakson. 1994. Pharmacological and biochemical demonstration of the role of cyclooxygenase 2 in inflammation and pain. *Proc. Natl. Acad. Sci. USA.* 91:12013–12017.
16. Hippisley-Cox, J., and C. Coupland. 2005. Risk of myocardial infarction in patients taking cyclo-oxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs: population-based nested case-control analysis. *BMJ.* 330:1366.
17. Luong, C., A. Miller, J. Barnett, J. Chow, C. Ramesha, and M. F. Browner. 1996. Flexibility of the NSAID binding site in the structure of human cyclooxygenase-2. *Nat. Struct. Biol.* 3:927–933.
18. Gierse, J. K., J. J. McDonald, S. D. Hauser, S. H. Rangwala, C. M. Koboldt, and K. Seibert. 1996. A single amino acid difference between cyclooxygenase-1 (COX-1) and -2 (COX-2) reverses the selectivity of COX-2 specific inhibitors. *J. Biol. Chem.* 271:15810–15814.
19. Guo, Q., L.-H. Wang, K.-H. Ruan, and R. J. Kulmacz. 1996. Role of Val^{509} in time-dependent inhibition of human prostaglandin H synthase-2 cyclooxygenase activity by isoform-selective agents. *J. Biol. Chem.* 271:19134–19139.
20. Wong, E., C. Bayly, H. L. Waterman, D. Riendeau, and J. A. Mancini. 1997. Conversion of prostaglandin G/H synthase-1 into an enzyme sensitive to PGHS-2 inhibitors by a double $His^{513} \rightarrow$ Arg and $Ile^{523} \rightarrow$ Val mutation. *J. Biol. Chem.* 272:9280–9286.
21. So, O.-Y., L. E. Scarafia, A. Y. Mak, O. H. Callan, and D. C. Swinney. 1998. The dynamics of prostaglandin H synthases. *J. Biol. Chem.* 273: 5801–5807.

22. Walker, M. C., R. G. Kurumbail, J. R. Kiefer, K. T. Moreland, C. M. Koboldt, P. C. Isakson, K. Seibert, and J. K. Gierse. 2001. A three-step kinetic mechanism for selective inhibition of cyclo-oxygenase-2 by diarylheterocyclic inhibitors. *Biochem. J.* 357:709–718.

23. Spencer, A. G., J. C. Thuresson, E. Otto, I. Song, T. Smith, D. L. DeWitt, R. M. Garavito, and W. L. Smith. 1999. The membrane binding domains of prostaglandin endoperoxide H synthases 1 and 2. *J. Biol. Chem.* 274:32936–32942.

24. DeWitt, D. L., and W. L. Smith. 1988. Primary structure of prostaglandin G/H synthase from sheep vesicular gland determined from the complementary DNA sequence. *Proc. Natl. Acad. Sci. USA.* 85:1412–1416.

25. Otto, J. C., and W. L. Smith. 1996. Photolabeling of prostaglandin endoperoxidase H synthase-1 with 3-trifluoro-3-(*m*-[^{125}i]iodophenyl)-diazirine as a probe of membrane association and the cyclooxygenase active site. *J. Biol. Chem.* 271:9906–9910.

26. Molnar, F., L. S. Norris, and K. Schulten. 2000. Simulated (un)binding of arachidonic acid in the cyclooxygenase site of prostaglandin H_2 synthase-1. *Progr. React. Kinet.* 25:263–298.

27. Nina, M., S. Bernèche, and B. Roux. 2000. Anchoring of a monotopic membrane protein: the binding of prostaglandin H_2 synthase-1 to the surface of a phospholipid bilayer. *Eur. Biophys. J.* 29:439–454.

28. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD—visual molecular dynamics. *J. Mol. Graph.* 14:33–38.

29. Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. 1994. Molecular Biology of the Cell. 3rd Ed. Garland Publishing, New York.

30. Roux, B., and K. Schulten. 2004. Computational studies of membrane channels. *Structure.* 12:1343–1351.

31. Berman, H., J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, and P. Bourne. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.

32. Kalé, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999. NAMD2: greater scalability for parallel molecular dynamics. *J. Comput. Phys.* 151:283–312.

33. A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.

34. Ryckaert, J.-P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23: 327–341.

35. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an *N*-log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

36. Deol, S. S., P. J. Bond, C. Domene, and M. S. Sansom. 2004. Lipid-protein interactions of integral membrane proteins: a comparative simulation study. *Biophys. J.* 87:3737–3749.

37. Isralewitz, B., M. Gao, and K. Schulten. 2001. Steered molecular dynamics and mechanical functions of proteins. *Curr. Opin. Struct. Biol.* 11:224–230.

38. Ulmschneider, M. B., M. S. P. Sansom, and A. Di Nola. 2005. Properties of integral membrane protein structures: derivation of an implicit membrane potential. *Proteins.* 59:252–265.

39. Kurumbail, R. G., A. M. Stevens, J. K. Gierse, J. J. McDonald, R. A. Stegeman, J. Y. Pak, D. Gildehaus, J. M. Miyashiro, T. D. Penning, K. Seibert, P. C. Isakson, and W. C. Stallings. 1996. Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents. *Nature.* 384:644–648.

40. Selinsky, B. S., K. Gupta, C. T. Sharkey, and P. J. Loll. 2001. Structural analysis of NSAID binding by prostaglandin H2 synthase: time-dependent and time-independent inhibitors elicit identical enzyme conformations. *Biochemistry.* 40:5172–5180.