Title

Solution Conformations and Dynamics of Substrate-Bound Cytochrome P450 MycG

Journal

Biochemistry, 56(21)

ISSN

0006-2960

Authors

Tietz, Drew R

Podust, Larissa M

Sherman, David H

et al.

Publication Date

2017-05-30

DOI

10.1021/acs.biochem.7b00291

Peer reviewed

Published in final edited form as:

Biochemistry. 2017 May 30; 56(21): 2701–2714. doi:10.1021/acs.biochem.7b00291.

Solution conformations and dynamics of substrate-bound cytochrome P450 MycG

Drew R. Tietz[a], **Larissa M. Podust**[c], **David H. Sherman**[d], and **Thomas C. Pochapsky**[a,b,*]

[a]Department of Chemistry, Brandeis University, MS 015, 415 South St., Waltham, MA 02453

[b]Rosenstiel Basic Medical Sciences Research Institute, Brandeis University, MS 015, 415 South St., Waltham, MA 02453

[c]Skaggs School of Pharmacy & Pharmaceutical Sciences, University of California, San Diego, CA 92093, USA

[d]Life Sciences Institute, University of Michigan, Ann Arbor, MI, 48109-2216

Abstract

MycG is a P450 monoxygenase that catalyzes the sequential hydroxylation and epoxidation of mycinamicin IV (M-IV), the last two steps in the biosynthesis of mycinamicin II, a macrolide antibiotic isolated from *M. griseorubida*. The crystal structure of MycG with M-IV bound was previously determined, but showed the bound substrate in an orientation that did not rationalize the observed regiochemistry of M-IV hydroxylation. NMR paramagnetic relaxation enhancements (PRE) gave evidence for an orientation of M-IV in the MycG active site more compatible with the observed chemistry, but substrate-induced changes in the enzyme structure were not characterized. We now describe the use of amide ^{1}H-^{15}N residual dipolar couplings (RDCs) as experimental restraints in solvated “soft annealing” molecular dynamics simulations to generate solution structural ensembles of M-IV-bound MycG. Chemical shift perturbations, hydrogen-deuterium exchange and ^{15}N relaxation behavior provide insight into dynamic and electronic perturbations in the MycG structure in response to M-IV binding. The solution and crystallographic structures are compared, and the possibility that the crystallographic orientation of bound M-IV represents an inhibitory mode is discussed.

Introduction

Macrolides are bioactive molecules containing a macrocyclic lactone ring, typically 12–16 atoms in size and substituted with one or more sugars, aza-sugars or other functional groups (1). Macrolides are bacteriostatic at low concentrations, but can become bacteriocidal at higher levels (2). Macrolide antibiotics bind reversibly to a site on the 50S subunit of the ribosome, thereby interfering with translocation and protein synthesis (3). The biosynthesis of macrolides involves successive Claissen condensations of malonyl and/or methylmalonyl

*To whom correspondence should be addressed: pochapsk@brandeis.edu.

Supplementary Information: Complete descriptions of plasmids and expression systems used, purification protocols, characterization of purified MycG and tables of chemical shift perturbation, H/D exchange and ^{15}N relaxation data used to generate Figure 3.

CoA subunits into a polyketide chain. These are directed by a multi-subunit polyketide synthase (PKS) complex, with each successive CoA addition to the polyketide providing the thioester linkage to the next PKS subunit (4, 5). When the appropriate chain size is reached, the lactone is closed by reaction of the free hydroxyl of the first subunit with the thioester of the last. Finishing steps in the biosynthesis include glycosidic linkage of sugars or aza-sugars to the macrolactone, and (often) oxidative elaboration of the macrolactone by cytochrome P450 monooxygenases. The regio- and stereoselectivity of these P450-catalyzed oxidations confer the bioactivity of the resulting macrolide, so the mechanistic details of how such oxidations take place are of interest for many reasons, including potential re-engineering to generate novel antibiotics (6, 7).

The mycinamicins are 16-member macrolides isolated from the actinomycete *Micromonospora griseorubida* (8). Two cytochrome P450s have been identified in the mycinamicin biosynthetic pathway, MycCI and MycG (2, 9). MycCI hydoxylates the C-21 methyl group of the macrolide prior to appending the mycinose precursor javose (see Scheme 1 for annotation of the mycinamicin carbon skeleton), while MycG catalyzes the last two steps leading from M-IV to M-II, a hydroxylation at C-14 (M-V) and epoxidation of the double bond between C-12 and C-13 yielding the final product M-II (Scheme 1). The reaction order (hydroxylation, then epoxidation) is obligate: Prior epoxidation of the double bond results in the dead-end product M-I (10). MycG is a member of a growing class of multifunctional P450 enzymes that catalyze several steps in a biosynthetic pathway while maintaining the requisite regio- and stereospecificity of the oxidation at each step (11). Such enzymes are attractive from a bioengineering standpoint, but their multi-functionality increases the difficulty of understanding the role a particular enzyme-substrate complex plays in directing catalysis. We have previously shown that solution nuclear magnetic resonance (NMR) data can be used to characterize conformational equilibria in cytochromes P450 as a function of specific conditions, thereby reducing the ambiguity regarding the role of a crystallographically observed conformation in enzyme function. For example, crystallographic structures of cytochrome $P450_{cam}$ (CYP101A1) rationalize the regio- and stereochemistry of the observed camphor hydroxylase activity (12), but do not account for the necessity of effector binding for substrate turnover. Using NMR, we identified an effector-induced *trans-cis* isomerization of a single Ile-Pro peptide bond that closes access to the active site and orients the substrate appropriately for hydroxylation (13). In subsequent work, we made use of residual dipolar couplings (RDCs) as restraints in solvated molecular dynamics (MD) simulations of CYP101A1 with (14) and without (15) substrate bound to identify conformational changes that take place upon substrate binding. These changes are not limited to the active site (which is largely collapsed in the absence of substrate) but are propagated throughout much of the enzyme structure. Using site-directed mutagenesis, we recently showed that these conformational changes are not artifacts of the methodology, but true representations of conformational selection upon substrate binding (16). Similar results have been noted recently for another P450 enzyme (17). Herein, we describe how solution NMR methods applied to MycG with substrate mycinamicin IV (M-IV) bound give insight into initial substrate binding conformations and dynamical changes, and help in the interpretation of our previously published crystallographic structures (10).

Experimental procedures

Protein expression and purification

Complete descriptions of expression plasmids and strains used for expression, media used for preparing isotope-enriched samples of MycG and protein purification methods are available in Supplementary Information.

NMR sample preparation

After purification, samples of MycG were stored in liquid nitrogen and thawed in a room temperature water bath immediately before spectroscopy. Samples were then exchanged into degassed and filtered NMR buffer (50 mM KPi, pH 7.4, 200 mM KCl, 10 % D_2O) by diluting from ~1 mL to ~15 mL, re-concentrating to ~350 μL with addition of M-IV to a final concentration of 1 mM Typical sample concentrations were 200–500 μM MycG. Samples were transferred to a susceptibility-matched NMR tube (Shigemi, Inc., Allison Park, PA) for spectroscopy.

NMR spectroscopy

All NMR spectra were acquired at 298 K on a Bruker Avance spectrometer (Brandeis University Landsman Research Facility) equipped with a 5 mm TXI cryoprobe and operating at 800.13, 201.20, and 81.086 MHz for ^{1}H, ^{13}C, and ^{15}N respectively. Standard TROSY-based pulse sequences were used to acquire 2D-^{1}H,^{15}N-HSQC, 2D–HNCO, 3D-HNCA, 3D-HN(CO)CA, 3D-HNCACB and ^{15}N-edited 3D-NOESY data sets with triply labeled (U-^{13}C, ^{15}N, ^{2}H) MycG samples. 2D-HSQC (U-^{15}N, or U-^{15}N,^{2}H-MycG) and 2D-HNCO (U-^{15}N, ^{13}C-Pro-MycG) experiments were obtained with 2048 (^{1}H) × 128 (^{15}N) complex points and 16 scans per t_1 increment. Triple resonance experiments were run with 2048 (^{1}H) × 96 (^{13}C) × 64 (^{15}N) complex points and 16 scans per t_1 increment, with the exception of the HNCACB which was run with 32 scans per t_1 increment. Sweep widths were 14367 Hz (^{1}H) and 2919 Hz (^{15}N) for all experiments, while ^{13}C spectral widths were 6438 Hz [HNCA and HN(CO)CA] and 15090 Hz for the HNCACB. The 3D, ^{1}H, ^{15}N NOESY-TROSY was acquired with 2048 × 40 × 128 complex points using 12818 Hz (^{1}H), 2919 Hz (^{15}N) and 12818 Hz (^{1}H) spectral widths respectively, using a 100 ms mixing time and a recycle delay of 1.5 seconds. All data sets were processed using Bruker Topspin 3.2 and analyzed using CCPN 2.3.1 or Sparky.

Titration of MycG with M-IV

A 115 mM stock of mycinamicin IV (M-IV) was prepared in d_6-dimethylsulfoxide (CIL). A primary consideration for the titration was to keep the overall DMSO concentrations in the protein sample to below 1%, while maintaining high volume precision for each added aliquot. This was achieved by adding 0.4 μL of 115 mM M-IV stock per titration point. 500 μL of 460 μM U-^{15}N MycG with R > 1.6 (A_{417}/A_{280}) was placed in a standard 5 mm NMR tube. 0.4 μL of 115 mM stock M-IV was transferred to each of 6 small glass reaction vials and tightly capped. A ^{1}H, ^{15}N TROSY-HSQC spectrum was acquired first for the substrate free MycG as described above. After each titration point a portion (~300 μL) was transferred via glass pipette from the NMR tube to the reaction vial containing the amount of M-IV

required for the next titration point. The protein solution was mixed thoroughly and returned to the NMR tube for acquisition of the next titration point. The procedure was repeated using DMSO stock with no M-IV to act as a control. A total of six titration points were obtained, with M-IV/MycG ratios ranging from 0.2 to 2 (Table 1), which adequately saturates the M-IV active site based on a 700 nM K_d (10).

Hydrogen-deuterium exchange measurements

Exchange reactions were initiated by concentrating U-^{15}N MycG to approximately 3 mM, then diluting 10-fold with deuterated buffer (50 mM KPi pH 7.4, 200 mM KCl, 99% D2O, (2 mM M-IV for substrate bound experiments)) immediately before transfer to an NMR tube for data acquisition. Control experiments showed that concentration of MycG to 3 mM, followed by 10-fold dilution yields HSQC spectra identical to an unprocessed sample. Lag times between exchange initiations were also kept to a minimum by locking, tuning, shimming and calibrating pulses on an identical sample (90 % D2O, 50 mM KPi pH 7.4, 200 mM KCl, 300 μM MycG) in the same NMR tube as the experiment to be performed. This method kept lag times to less than two minutes. A continuous series of TROSY-HSQC experiments were recorded over the course of 20 hours. Experiments were performed with 2048 (^{1}H) and 96 (^{15}N) complex points. Spectral widths were 14367 Hz (^{1}H) and 2919 Hz for (^{15}N). For each t1 increment 16 transients were recorded resulting in a total acquisition time of 30 minutes per TROSY-HSQC.

^{15}N T_1 and T_2 measurements

A sample containing 500 μM ^{15}N MycG and 1 mM M-IV was prepared as described above. ^{15}N T_1 experiments were recorded using a standard inversion-recovery sequence with 2048 (^{1}H) × 128 (^{15}N) complex points, 32 scans per t1 increment and 2.5 second recycle delay. A τ delay list of 100, 200, 400, 800, 1600, and 2400 ms was used and acquired in an interleaved fashion to compensate for any sample changes or spectrometer drift over the course of the experiment. ^{15}N T_2 datasets were obtained using a standard CPMG sequence with 2048 (^{1}H) × 128 (^{15}N) complex points, 32 scans per t1 increment and 1.0 second recycle delay. A variable delay list of 0, 16, 32, 48 and 64 ms was used in an interleaved manner. Data processing was performed using Bruker Topspin 3.2 and relaxation analysis was done using CCPN 2.3.1.

RDC measurements

Residual dipolar couplings were measured using TROSY/semi-TROSY datasets obtained as described by Weigelt (18). Interleaved data sets were acquired with 2048 (^{1}H) × 512 (^{15}N) complex points with 14367 Hz (^{1}H) and 3243 Hz (^{15}N) spectral widths. For each t1 increment 64 transients were recorded. Prior to Fourier transformation the interleaved data sets were separated yielding two data sets each with 2048 (^{1}H) × 256 (^{15}N) complex points. One level of linear prediction was applied in the ^{15}N dimension of all data sets, followed by application of a 90°-shifted square sine bell, and zero filling to 1024 points. The ^{1}H dimension was treated with the same window function and zero filled to 4096 points. Fourier transformation of each data set pair resulted in two spectra offset in the ^{1}H dimension by $^1J_{NH}$. The $^1J_{NH}$ offsets were measured in Hz and subtracted

($^1J_{NH}$(aligned)-$^1J_{NH}$(unaligned)) to determine residual dipolar coupling contributions to the aligned $^1J_{NH}$ measurements.

Alignment media preparation for residual dipolar coupling (RDC) measurements, $C_{12}E_5$

175 μL of a 10% (v/v) solution of pentaethylene glycol monododecyl ether ($C_{12}E_5$, 98+% Fluka, Switzerland) was prepared in degassed NMR buffer (50 mM KPi (pH 7.4) and 200 mM KCl in 80:20 H_2O:D_2O) and vortexed until homogeneous and clear. Anhydrous n-hexanol (Aldrich) was added in 0.5 μL aliquots and briefly vortexed after each addition. The originally clear solution became milky and viscous upon first addition of hexanol. After 6.5 μL of hexanol had been added, the solution became opalescent and only slightly more viscous than water. In practice it was found to be optimal after addition of the hexanol aliquot that turned the solution opalescent, one subsequent addition 0.5 μL hexanol aliquot was added to bring the mixture away from boundaries of the phase diagram (19). The final $C_{12}E_5$/hexanol molar ratio was 0.85. Care was taken to keep the temperature of the media within a range close to the planned NMR acquisition temperature of 25°C during preparation, as liquid crystal phase transitions are sensitive to both chemical composition and temperature. An aliquot of U-^{15}N MycG (500 μL, 500 μM, 1 mM M-IV) was thawed by submerging in a room temperature water bath. The thawed protein was concentrated to ~1.0 mM using a Mllipore concentrator with a 10 kDa molecular weight cutoff. A 1:1 (v/v) ratio of 1 mM MycG solution and the prepared $C_{12}E_5$/hexanol solution (175 μL of each) were mixed by pipetting, then very lightly and briefly vortexed. The sample retained its transparency and did not appear significantly different than a standard MycG NMR sample without alignment media. The entire prepared sample (350 μL) was transferred to a susceptibility matched NMR tube (Shigemi, Inc., Allison Park, PA). Alignment was confirmed by observing quadrupolar splitting of the ^{2}H buffer signal. The splitting was measured to be 28.2 Hz after 30 minutes in the magnet and did not change after the initial measurement.

Alignment media preparation, pf1 phage

100 μL of a 50 +/−4 mg/mL *pf1* bacteriophage solution (Asla Bioteck Ltd., Riga, Lativa) was added to 400 μL of 625 μM U-^{15}N MycG in M-IV-containing NMR buffer. The mixture containing a final *pf1* concentration of 10 mg/mL, 1 mM M-IV and 500 μM MycG was gently vortexed to homogeneity. The sample was transferred to a susceptibility matched NMR tube (Shigemi, Inc., Allison Park, PA). A quadrupolar ^{2}H splitting of 9 Hz of the water line was observed and remained stable throughout the experiment.

RDC-restrained molecular dynamics simulations

All simulations were performed using the AMBER 15.0 package operating in parallel (.mpi) mode on the XSEDE node Comet (University of California San Diego). Initial coordinates for simulations were obtained from the MycG-M-IV crystallographic structure (PDB code 2Y98) (10). Substrate M-IV parameters were obtained using coordinate and connectivity data from the crystallographic structure. Hydrogens were added to M-IV using PyMOL version 1.3 and saved as a PDB file, which was subsequently converted to mol2 format without assigned charges. Charges for each atom were calculated and assigned with the *antechamber* module of *ambertools 1.5* Amber SQM version 1.4 utilizing a semi-empirical

(AM1-BCC) method. The *parmchk* utility was used to parameterize the force field using the general Amber force field (GAFF) as a starting point. Ferric (Fe^{3+}) heme topology and force field parameters were those previously determined and deposited in the AMBER parameter database (20).

To prepare the polypeptide for simulations, waters of solvation were stripped from the structure 2Y98, and cysteine and histidine residues renamed to accurately reflect the appropriate pronation states. Specifically, Cys 346 was assigned to an unprotonated state (CYM) for covalent linkage to the heme iron and all histidines were assigned to be singly protonated at the Nε position. The modified PDB file was then loaded into *xleap* and recombined with the previously prepared heme and M-IV models using the 2Y98 coordinates for all atoms. A covalent linkage was created from the heme iron to Cys 346 with a Fe-S bond length of 2.45 angstroms (21). The water distally coordinated to the heme iron was retained based on evidence found in the 2Y98 (MycG-M-IV) and 2YGX (MycG) crystallographic structures. The complex was solvated within a cubic periodic box with 90 Å edges, leaving a minimum 15 Å distance from the periodic boundary to the protein surface. A total of 22,098 TIP3P water molecules, 26 K^+ and 13 Cl^- ions were added to solvate and neutralize the system, generating a 0.2 M KCl concentration that matches the ionic strength of buffer used in RDC measurements.

For all simulations Amber FF14SB force field was used for MycG and a modified GAFF force field parameterized with *antechamber* (AmberTools) was used for M-IV. Heme parameters were those obtained from the Amber parameter database. Prior to application of RDC restraints the structure was equilibrated as follows. An initial minimization was performed over 4000 steps while holding the protein fixed with 500 kcal mol^{-1} $Å^{-2}$ harmonic restraints followed by 8000 steps of minimization with the positional restraints released. Heating to 300 K was accomplished over 100 ps in three stages (0–40 K, 40–100 K, 100 –300 K) under constant volume conditions using a Langevin thermostat with a 1.0 ps time constant and a collision frequency of 2.0 ps^{-1}. After heating, an additional 200 ps of equilibration at 300 K was performed. Finally, a 1 ns constant pressure equilibration to 1 atm was performed using a pressure relaxation time of 2.0 ps. Time steps of 2 fs were used throughout with bonds involving hydrogen (on the protein, not M-IV or heme) constrained to their equilibrium lengths using the SHAKE algorithm. We have previously described the use of paramagnetic relaxation enhancements (PRE) to obtain relative distance restraints between individual hydrogen atoms on M-IV and the heme iron (10). These restraints were applied during equilibration in order to generate a reasonable starting structure for the complex. The PRE-based restraints on M-IV were maintained throughout the RDC-restrained simulations described below, at a reduced penalty so as not to distort the effects of RDCs on the overall conformations of the polypeptide.

After heating and equilibration, an initial minimization with fixed atom positions and RDC restraints active was performed to obtain preliminary alignment tensors for the two sets of RDC data. RDC restraints were then introduced gradually with protein atoms free to move, with penalties increased from 0.001 to 0.01 kcal $mol^{-1}(\delta D)^{-2}$ over several nanoseconds of simulation at 300 K, where δD is the difference between the observed and calculated RDC exceeding the +/− 3 Hz uncertainty of measurement for a given restraint. The final penalty

value used for production runs (0.01 kcal $mol^{-1}(\delta D)^{-2}$) was found in our previous work to impose a reasonable restraint set without distortion of local geometry (14). Runs were halted every 0.25 nsec in order to detect and correct any significant violations of RDC restraints, either due to misassignment or incorrect measurement. The alignment tensors for each set of RDC data were interactively optimized during this process as well. Optimization was continued until RDC violations reached a plateau and simulation densities maintained a constant physically realistic average. These conditions were reached after ~ 8.2 ns of simulation. Production runs were performed for another 6.8 ns, with 10 structures showing the lowest energy for RDC violations selected for further analysis using the AmberTools program *cpptraj*.

Results

Sequential assignment of backbone resonances of MycG-M-IV

MycG is now the second cytochrome P450 for which comprehensive backbone ^{1}H, ^{13}C and ^{15}N assignments have been made (CYP101A1 being the first) (22). Obtaining sequential resonance assignments for a cytochrome P450 presents significant challenges. The size of the enzyme (typically ~ 45 – 47 kDa for soluble P450s) demands that the protein be fully deuterated in order for multidimensional NMR methods to succeed. Whereas many perdeuterated proteins can be unfolded and then refolded in protonated buffers to recover the signals of slow-exchange amide protons, denaturation/refolding efforts with P450s are low in yield, making slow-exchange amides inaccessible to ^{15}N-attached ^{1}H-detected methods. Even NH positions that exchange sufficiently rapidly might not be assigned if they are in a region of otherwise slow exchange, making sequential correlations impossible. Compounding this problem is the fact that there are 27 prolines (Pro) in the MycG sequence, resulting in breaks in the backbone amide NH-based sequential assignment methods to which standard triple-resonance NMR experiments lend themselves.

An additional roadblock for NMR of cytochromes P450 is presented by heme-based paramagnetic relaxation. The heme iron is typically low spin Fe^{+3} (S=1/2) in the absence of substrate, and there is often some degree of spin-state change to S=5/2 (hgh spin) accompanying substrate binding. Both states are sufficiently paramagnetic to induce efficient relaxation in nearby spins: We have found that ^{1}H spins within ~14 Å of the heme iron are affected to some degree in high-spin samples of CYP101A1. For MycG, the conversion to high spin upon binding of MIV is ~15%, so paramagnetic relaxation effects are less prominent than in CYP101A1, which shows nearly complete conversion to the high-spin form upon substrate binding. We were therefore able to make assignments of many resonances in the vicinity of the active site that could not be obtained for CYP101A1 in the substrate bound form. Still, we made extensive use of type-selective ^{15}N labeled samples of MycG to provide multiple unambiguous starting points for the backbone-based (HNCA, HN(CO)CA and HNCACB) assignment process. Samples were selectively ^{15}N-labeled at Val (34), Ile (10), Leu (43), Phe (10), Tyr (9) and Arg (42), with the number of residues of each type in the MycG sequence shown in parentheses. To overcome the presence of 27 Pro residues in MycG, a sample was prepared uniformly labeled with ^{15}N and selectively labeled at Pro with ^{13}C. A two-dimensional ^{1}H, ^{15}N HNCO spectrum of this sample shows

prominent correlations for NH groups following the $^{13}C=O$ of proline residues, and was used to identify amide NH correlations immediately following Pro in the sequence. While some scrambling was observed in this sample (^{13}C from proline seen in other residues), relative signal intensities were usually sufficient to distinguish true Pro-X correlations from scrambled sites. The 3D ^{15}N-edited NOESY spectrum of perdeuterated ^{15}N-labeled MycG also proved valuable for confirming sequential NH assignments in helical regions as well as cross-strand assignments in β-sheets. As expected, NOEs were weak to unobservable in protonated MycG samples: The deuteration of nonexchangable sites was thus critical not only for detecting triple-resonance correlations but for detection of ^{1}H NOEs between exchangeable sites.

The assignments described here for MycG are somewhat more complete (~82% of non-proline backbone NH groups) than those for CYP101A1 (~75%). It appears that amide $^{2}H/^{1}H$ exchange in MycG during purification in protonated buffers is more complete than in CYP101A1 for perdeuterated samples. The substrate of MycG, M-IV, is both larger and more hydrophilic than camphor, the substrate of CYP101A1. As discussed below, the active site of MycG is larger than that of CYP101A1, and does not exclude water from the interior of the protein in the MycG-M-IV complex.

Solution structural ensembles from RDC-restrained soft annealing of M-IV-bound MycG

We have previously described the application of RDC-restrained molecular dynamics simulations to characterizing solution conformational ensembles for CYP101A1 (14, 15). The crystallographic structure 2Y98 was used as the starting point for the simulations as described above. The substrate M-IV was positioned in the active site as observed crystallographically. After an initial minimization of 4000 steps, the temperature of the simulation was raised slowly under constant volume conditions to 300 K, followed by 1 ns of equilibration under constant pressure (1 atm) conditions. In the course of heating and equilibration, the maximum distances of specific M-IV hydrogens relative to the heme iron were maintained using distance restraints obtained by paramagnetic relaxation enhancement (PRE) as described previously (10). No other experimental restraints were applied until the density of the simulation equilibrated at 1.014 kg/L. At this point, two sets of RDC restraints on backbone amide N-H bond vectors were introduced, one measured in the presence of *pf1* phage and the other in a nematic liquid crystalline phase ($C_{12}E_5$). Preliminary sets included 268 restraints from *pf1* and 248 restraints from $C_{12}E_5$. As restraint violation energies were gradually increased, incorrect assignments and/or inaccurate measurements were identified and either removed or corrected. This iterative “soft annealing” process resulted in a refined set of 252 RDC restraints from *pf1* and 222 RDC restraints from $C_{12}E_5$. We found that repeat violations were minimal after a total of 8.2 ns of simulation (including heating and constant volume equilibration). After this point, seven coordinate sets showing the lowest total restraint violation energies were extracted from the dynamics tracks. These sets occurred with a frequency of ~ 3/ns, so the runs continued until a total of 15 ns were accumulated. A superposition of the backbone structures of the selected coordinates are shown in Figure 1. A representative structure from the ensemble, obtained at 13.345 ns from the production run, was chosen as REP1 for further evaluation (Figure 2).

Chemical shift perturbations upon binding of M-IV to MycG

The ^{1}H,^{15}N TROSY-HSQC spectrum provides a rapid and residue-specific measurement of perturbations in the MycG structure upon substrate binding. Amide NH correlations are sensitive to changes in the local electronic environment, and perturbations are often seen in hinge regions between secondary structural elements, where local hydrogen bonds might be affected by changes in relative orientations of secondary structures. Upon binding of M-IV, resonances throughout the MycG structure exhibit perturbations. A complete list of the identified perturbations is provided in the Supplementary Information (Table S1), and the spatial distribution of those perturbations in the MycG structure is shown in Figure 3. “Slow exchange” indicates that discrete resonances are observed only in the absence of substrate and upon substrate saturation, while intermediate to fast exchange indicates that discrete but broadened resonances are observed for a particular residue at each titration point.

It is notable that many of the M-IV-induced perturbations exhibit slow-exchange behavior, indicating that some conformational changes associated with substrate binding occur on the ms time scale. Most of those correlations that exhibit slow exchange behavior are perturbed by > 20 Hz in the ^{1}H dimension at 298K (Figures 3 and 4). From estimated line widths (typically ~35 Hz at half-height for M-IV bound MycG), slow exchange behavior predicts state lifetimes > 10 ms and exchange rates slower than 50 s^{-1}. This behavior is not uniform across all perturbed resonances: In some cases, intermediate exchange behavior is observed for perturbations greater than 20 Hz, suggesting locally faster exchange between multiple conformations rather than a high-barrier shift between two conformational ensemble minima. For example, Tyr 39, located on a loop between the β1 strands, moves 116 Hz in the ^{1}H dimension in the course of the titration, yet shows intermediate exchange behavior.

Hydrogen-deuterium (H/D) exchange measurements

H/D exchange as measured by NMR provides a measure of solvent accessibility into the enzyme structure on a minute-to-hour time scale. While this time scale is not particularly relevant for mechanistic considerations, it provides insight into the relative mobility of different structural features, and is suggestive of which regions of the enzyme are likely to be involved in large-amplitude motions. In the case of MycG, many amide sites exchange too rapidly to be characterized by NMR, while others exchange too slowly to detect any change. However, a comparison between H/D exchange in substrate-free and M-IV-bound MycG shows that M-IV binding has the counterintuitive effect of increasing exchange rates at several sites, mostly in hinge regions (see Figure 3).

^{15}N relaxation behavior of M-IV-bound MycG

^{15}N relaxation behavior can be related to internal motions of a macromolecule because of the dependence of relaxation efficiency on the local spectral density, $J(\omega)$. The spectral density is a measure of the contribution of a particular radial frequency ω to the electromagnetic “white noise” in the molecular environment. Nuclear spin relaxation is most efficiently induced by fluctuations occurring at the frequency of the relevant spin transition, and spectral densities can be extracted by careful analysis of nuclear spin relaxation rates. In turn, the spectral densities provide insight into local dynamics. In the simplest model, that of Lipari and Szabo (23), the spectral density can be related to three parameters, the overall

molecular correlation time (τ_m), the observed correlation time (τ) and an order parameter, S, which is a measure of the degree to which local motions are constrained to the overall motion of the molecule (Eq. 1).

$$J(\omega)=\frac{2}{5}\left[\frac{S^2\tau_m}{1+\omega^2\tau_m^2}+\frac{(1-S^2)\tau}{1+\omega^2\tau^2}\right] \quad (1)$$

In turn, τ is related to τ_m by:

$$\tau^{-1}=\tau_m^{-1}+\tau_e^{-1} \quad (2)$$

where τ_e is a measure of local fluctuations. However, an even simpler analysis of ^{15}N relaxation behavior is possible: By consideration of the terms common to the equations relating observed T_1 and T_2 to spectral densities, the ratios of the observed relaxation times for a given residue provides a measure of information regarding local versus overall molecular dynamics (24). Large amplitude motions on the ns time scale will increase T_2 time while decreasing T_1 times, leading to T_1/T_2 decreasing as the frequency of internal motions increase. Following the same trend, as motions move toward lower frequencies approaching the millisecond time scale, T_2 will decrease while T_1 increases, resulting in a T_1/T_2 ratio that increases. Residues with T_1/T_2 ratios more than one standard deviation from the mean are interpreted to have significant contribution of internal motion (τ_e) to the effective correlation time. The time scale of these internal motions can be classified as slow (milliseconds), or fast (nanoseconds to picoseconds) depending on whether or the T_1/T_2 ratios are above or below the mean, respectively.

Table 2 lists those residues for which the ratio lie more than one standard deviation above or below the mean value, with secondary structural features noted. Figure 3 shows the spatial distribution of those T_1/T_2 ratios for ^{15}N resonances listed in Table 2.

Discussion

Comparison of crystallographic and RDC-derived structures and substrate-binding mode(s)

The primary effect of RDC restraints is to orient N-H bond vectors so as to minimize restraint energies relative to the principal components of an alignment tensor. In the case of MycG, two independent alignment tensors were used, one for RDCs measured by phage alignment and one for those measured in the $C_{12}E_5$ medium. While each individual RDC restraint is weak relative to other force field components such as electrostatics, van der Waals interactions, bond lengths and angles, the presence of multiple RDC restraints within a single secondary structural feature (e.g., an α-helix or β-sheet) exerts “pressure” on that element to orient so as to minimize violations. As such, the orientations of secondary structural elements in the molecular frame are typically more accurate than what is obtained using localized NMR restraints such as nuclear Overhauser effects and chemical shifts.

The crystallographic structure of M-IV-bound MycG (2Y98) provided the starting point for the solution structure calculations, and is the most relevant structure for comparison of the crystal and solution. The superposition of the two structures (REP1 and 2Y98) is shown in Figure 5. The backbone RMS deviation of the superposition is 1.49 Å. By far the most significant differences between the structures are in the positions of the F and G helices and the F-G loop: The crystallographic structure packs these features close to bound substrate. The more relaxed positions of the F-G loop and G helix in REP1 suggest that, in the absence of crystal packing constraints, these features sample a larger conformational space, opening the active site considerably. As part of this opening, the β-rich region (containing the β1, β3 and β4 sheets) appears to move as a unit to accommodate the modified substrate orientation described below.

We used the 3V (Voss Volume Voxilator) Web application to compare protein and active site volumes (25). The 2Y98 structure yielded a protein volume of 64 335 $Å^3$ and surface area of 11 996 $Å^2$. The vacated active site (with M-IV removed) exhibited as two separate voids, with volumes of 342 $Å^3$ and 245 $Å^2$ and internal surface areas of 418 $Å^3$ and 304 $Å^2$ respectively. The effective radius of each void was 4.56 Å and 4.1 Å, respectively. The larger void is located beneath the G helix and F-G loop, while the smaller one is located between the I helix and Phe 286, a primary substrate contact. These voids are shown in Figure 6.

As we observed in our comparisons of CYP101A1 crystallographic and RDC-derived structures, the volume of RDC-derived structure of MycG is larger than the crystallographic structure: 66929 $Å^3$ with a surface area of 12245 $Å^2$. Unlike the crystallographic structure, the active site for the RDC-derived structure exhibits as a single void with a volume of 977 $Å^3$ and internal surface area of 595 $Å^2$, and an effective radius of 4.93 Å (Figure 7). The M-IV molecule has a volume of 990 $Å^3$ and surface area of 625 $Å^2$, consistent with the significant solvent exposure of MycG-bound M-IV in both structures.

The orientation of M-IV in the active sites of the two structures differ significantly (Figure 8). Both show an orientation of the M-IV molecule with the mycinose methoxy groups close to the heme plane, but the long axis of the MIV molecule is roughly perpendicular to the heme plane in 2Y98 (“mycinose in-desosamine out”), while PRE restraints place the macrolactone ring in REP1 in an orientation more parallel to the heme, moving the desosamine ring into van der Waals contact with the B’-C loop, placing the side chains of Glu 77 and Lys 80 near the desosamine, with main chain contacts to the desosamine provided by Gly 81. Other substrate contacts in the REP1 active site include Val 79 (macrolide) and Leu 84 in the B’-C loop (mycinose), Phe 168 and Ser 170 in the F-G loop (macrolide), I helix residues Ile 229, Val 233 and Ser 238 (macrolide), Leu 386 and Leu 387 in the β5 turn (macrolide and mycinose) and Phe 286 near the N-terminal end of the β3 strand (mycinose) (Figure 9).

In 2Y98, Phe 168 and Ser 170 in the F-G loop also contact the macrolide, but Arg 75 replaces Glu 77 as the primary desosamine contact in the B’-C loop. Ala 234 (I helix) contacts the mycinose, but the macrolide ring moves away from the β5 turn, so that neither Leu 386 nor Leu 387 make significant substrate contacts. The desosamine moves away from the B-B’ loop, and instead packs against the G helix, with Met 179 being a primary

desosamine contact. This is markedly different from REP1, in which the G helix moves away from the desosamine: Inspection suggests that much of the extra volume both of the overall REP1 structure and the active site channel is due to this displacement of the G helix (vide supra), with the volume between the desosamine and the G helix being fully hydrated. Other evidence for compaction in the 2Y98 active site is the presence of a well-developed B" helix between Gly 82 and Met 86 in 2Y98, which while roughly helical is less ordered in REP1.

Both REP1 and 2Y98 differ from our earlier modeling obtained using only Fe-(M-IV) PRE restraints (10). Without experimental restraints on the polypeptide, we saw a closer approach of the hydroxylation site (C-14) to the heme iron than is seen in REP1 (6.8 Å versus 9.1 Å), although the orientations of M-IV in the PRE-derived model and REP1 are similar. Another important difference between the PRE-derived model and REP1 is the use of explicit solvent in our current calculations. In REP1, three water molecules (including the heme ligand) are found in the active site within 7 Å of the heme iron, forming "spacers" between the macrolide and the heme, whereas the lack of explicit solvent in earlier calculations permitted the closer approach of the hydroxylation site C-14 to the iron in response to the PRE restraints (no close-approach penalty was applied to PRE restraints in either simulation). There are four water molecules, including the heme ligand, between M-IV and the heme in 2Y98, with a slightly longer distance between C-14 and the heme iron than in REP1 (9.2 versus 9.1 Å).

The relatively small fraction of high spin (~15%) formed upon M-IV saturation of the MycG active site suggests that more complete desolvation of the active site may be necessary to form a productive (high-spin) complex. Typically, formation of the high-spin form of a P450 is required for the first electron transfer to occur, as the reduction potential of the low-spin heme iron is mismatched with that of most physiological reductants (26). Such a mechanism would allow reduction and subsequent O_2 binding to occur only if M-IV binds so as to displace active site waters. This might also rationalize the obligate order (hydroxylation before epoxidation) if the more polar M-I results in a tendency to retain waters of hydration upon binding. The number of water molecules in the active site remains relatively constant after equilibration throughout our simulations. Given the enforced presence of the axial water ligand of the heme iron, it is possible that a more completely desolvated active site might never be accessed using our current configuration. We are currently performing simulations without the ligating water to see if a desolvated state is ever accessed with RDC and PRE restraints active.

Chemical shift and dynamic perturbations in MycG upon binding of M-IV

The global distribution of chemical shift perturbations observed upon binding of M-IV to MycG confirms that substrate binding involves many of the secondary structural features of the enzyme, not just the active site residues. Amide chemical shift perturbations reflect local electronic changes (hydrogen bond lengths and angles) so it is not surprising that many of the perturbations occur at hinge regions or at contacts between secondary structures. There are striking similarities between substrate-dependent perturbations observed for MycG and what was observed previously with CYP101 A1. Despite the differences in size of substrate

and active sites, many of the same regions are perturbed upon substrate binding in both enzymes. Besides residues adjacent to the active sites, the β-rich regions, including the β1, β3 and β4 sheets, A, B and K' helices are strongly perturbed in both MycG and CYP101A1 (15). The junction between the C and D helices, near the top of the structure as viewed in Figure 4, are also affected in both enzymes, as are the junctions of the I helix at both ends. This suggests that the same movements, involving many of the same structural features, occur in both enzymes upon substrate binding.

There are differences between the responses of the two enzymes to substrate binding as well. The N-terminal of the G helix is more strongly perturbed in MycG than in CYP101A1, perhaps reflecting the fact that there is no interaction between substrate and the G-helix in CYP101A1 as there is in MycG, or that the substrate being so much larger in MycG that larger motions are required to accommodate M-IV than the small substrate camphor, or, as discussed below, the crystallographic orientation of M-IV is populated to some extent in solution. The residues in the I helix "kink" directly over the heme are less responsive in MycG than in CYP101A1. This may be due to the observation that only 15% of the M-IV-bound MycG is high-spin, while camphor binding to CYP101A1 induces nearly complete conversion to the high-spin form. Perturbations near the heme may reflect the changes that occur upon high spin formation.

Local dynamics and structure

Given the limited number of residues for which D/H exchange could be measured, it is difficult to draw conclusions regarding the effect of substrate binding on specific structural features. However, it is worth noting that in all but one case, D/H exchange was faster in the substrate bound enzyme than in the substrate-free form. This observation is somewhat counterintuitive, given that enzyme structures are typically stabilized in the presence of substrate. It is possible that substrate binding introduces large-scale normal vibrational modes that are inactive in the substrate-free form.

The ^{15}N relaxation data are somewhat more informative. Not surprisingly, those regions of the enzyme that show ns-ps dynamics (^{15}N T_1/T_2 ratios more than 1 standard deviation below the mean) tend to be found in loop regions between secondary structures external from the active site. In particular, the N-terminal residues sequentially prior to the A helix show fairly uniform dynamics in this regime, as does the loop between the β1 strands, as well as the B'-C, D-E, E-F, H-I and J-K loops. On the other hand, residues showing ms dynamic contributions (^{15}N T_1/T_2 ratios more than 1 standard deviation above the mean) tend to be clustered closer to the active site, consistent with the observed slow exchange behavior of many chemical shift changes associated with substrate binding.

Does the crystallographic M-IV binding mode represent an inhibitory complex?

It is now becoming evident that many P450 enzymes are capable of binding multiple substrates/inhibitor molecules, or else to bind substrates or inhibitors in positions that seem to be not conducive to the observed regiochemistry of substrate oxidation. Besides the present example provided by MycG, a recent crystallographic structure of CYP2D6 shows inhibitor quinidine bound in a cleft between the F and G helices relatively remote from the

heme (27), while PREs measured by solution NMR suggest that quinidine can bind in a more “substrate-like” orientation close to the heme iron in the CYP2D6 active site (Pochapsky et al, unpublished results). The crystallographic structure of CYP2C5 shows evidence for two orientations of substrate, also placed in a cleft between the F, F’, F” and G helices (28). $P450_{nor}$ (nitric oxide reductase) directly binds cofactor NADH in a cleft formed by the B’, F and G helices (29). Monk et al. reported a secondary substrate binding site within the membrane-binding domain of CYP51, which includes portions of the B’, F’ helices and the β1 and β3 sheets (30). CYP3A4 has long been known to bind multiple substrate molecules simultaneously (31, 32), and even CYP101A1 has been suggested to have a “vestibule” secondary camphor binding site near the F-G loop in that enzyme (33).

Substrate inhibition is also observed with a variety of P450 enzymes. Arachadonic acid is both a substrate and inhibitor of the CYP1A series (34), and carbamazepine shows substrate inhibition of CYP3A4 in membranes and in vitro (35). A kinetic profile of aniline metabolism by CYP2E1 was best fit by a two-site binding model, in which a high affinity site (K_d = 30 μM) is occupied first, followed by second-site binding (K_d = 1100 μM) that results in a threefold increase in turnover rate (36). In biosynthetic P450 enzymes, substrate inhibition has been observed by kinetic analysis of PikC (37), EryK (38) and PimD (39). M-IV inhibition of MycG was shown indirectly, as a considerable drop in the NADPH consumption rate at high M-IV concentration (e.g. 100–500 μM) compared to low M-IV concentration (e.g. 10 μM), clarifying a controversy between high M-IV binding affinity and low product/substrate ratio observed at 500 μM M-IV (2). These observations raise an intriguing question: Does the crystallographic orientation of M-IV in the MycG active site represent an inhibitory binding mode? We note that the F and G helices are more tightly packed with respect to both substrate and the remainder of the enzyme in the crystal than in the solution structure. While it is possible that this tight packing is in response to crystal packing restraints, it is equally likely that this binding mode represents a populated solution conformer that is more susceptible to crystallization than the open forms dominating the solution ensemble. Furthermore, given the remoteness of much of the M-IV molecule from the heme in this orientation, population of this conformer would not be detected readily using PRE restraints. In Figure 10, the crystallographic orientation of M-IV in MycG is compared with that of an inhibitor, quinidine, bound to CYP2D6, a human P450 involved in drug and steroid metabolism. Although the tilt of the F and G helices relative to the active site differ between the two enzymes, the helices maintain the same orientations with respect to each other, and the similarity in substrate (M-IV) and inhibitor (quinidine) binding modes is evident. The inhibitory behavior of quinidine (and substrate inhibition by M-IV) could result from a “locking” of functionally important motions of the F-G helices required for accessing the active site and/or catalytically important movements. Furthermore, if the dead-end epoxidation product M-I (Scheme I) binds tightly in the crystallographic orientation, this could explain the obligate order of MycG-catalyzed chemistry (first hydroxylation at C14, then epoxidation of the C12-C13 double bond).

Further support for this possibility is provided by the observation that NH correlations for residues in the F-G loop and F and G helices are perturbed upon M-IV binding (Fig. 2). The solution structure in Figure 2 shows no direct interaction between M-IV and the F-G loop, and as noted above, the volume between the F-G loop and M-IV in the solution structure is

fully solvated but otherwise vacant. If the crystallographic orientation of M-IV is partially occupied in solution, this could explain these observations, as well as the fact that the slow exchange regime for chemical shift perturbations varies depending upon the residue under observation.

Structure and NMR Assignment Deposition—Resonance assignments for substrate-bound MycG have been deposited in the BMRB with accession number 26750. Coordinates for the solution structure REP1 of MIV-bound MycG have been deposited in the RCSB PDB with the identifier 5UHU, along with PRE and RDC restraints and Amber-format parameter/topology and restart files.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

TCP and DRT thank Dr. Francesco Pontiggia for help in setting up calculations on the Brandeis University high-performance cluster and XSEDE nodes, and Dr. Eliana Asciutto (CONICET, Buenos Aires, Argentina) for helpful advice and discussion. This work was partially supported by the NIH (R01 GM44191, T.C.P.) and R01 GM078553 (D.H.S. and L.M.P.). TCP acknowledges a National Science Foundation starter grant of XSEDE access (MCB160203).

References

1. Satoi S, Muto N, Hayashi M, Fujii T, Otani M. Mycinamicins, new macrolide antibiotics. 1. Taxonomy, production, isolation, characterization and properties. J. Antibiot. 1980; 33:364–376. [PubMed: 7410205]
2. Anzai Y, Li SY, Chaulagain MR, Kinoshita K, Kato F, Montgomery J, Sherman DH. Functional analysis of MycCI and MycG, cytochrome P450 enzymes involved in biosynthesis of mycinamicin macrolide antibiotics. Chem. Biol. 2008; 15:950–959. [PubMed: 18804032]
3. Leclercq R, Courvalin P. Bacterial resistance to macrolide, lincosamide, and streptogramin antibiotics by target modification. Antimicrob. Agents Chemother. 1991; 35:1267–1272. [PubMed: 1929280]
4. Aparicio JF, Mendes MV, Anton N, Recio E, Martin JF. Polyene macrolide antibiotic biosynthesis. Curr. Med. Chem. 2004; 11:1645–1656. [PubMed: 15180569]
5. Aparicio JF, Caffrey P, Gil JA, Zotchev SB. Polyene antibiotic biosynthesis gene clusters. App. Microbiol. Biotechn. 2003; 61:179–188.
6. Zhang W, Liu Y, Yan JY, Cao SN, Bai FL, Yang Y, Huang SH, Yao LS, Anzai Y, Kato F, Podust LM, Sherman DH, Li SY. New reactions and products resulting from alternative interactions between the P450 enzyme and redox partners. J. Am. Chem. Soc. 2014; 136:3640–3646. [PubMed: 24521145]
7. Li SY, Chaulagain MR, Knauff AR, Podust LM, Montgomery J, Sherman DH. Selective oxidation of carbolide C-H bonds by an engineered macrolide P450 mono-oxygenase. Proc. Natl. Acad. Sci. USA. 2009; 106:18469–18474. [PubMed: 19841270]
8. Kinoshita K, Takenaka S, Suzuki H, Morohoshi T, Hayashi M. Mycinamicins, New macrolide antibiotics .13. Isolation and structures of novel fermentation products from Micromonospora griseorubida (FERM BP-705). J. Antibiot. 1992; 45:1–9. [PubMed: 1548179]
9. DeMars MD, Sheng F, Park SR, Lowell AN, Podust LM, Sherman DH. Biochemical and structural characterization of MycCI, a versatile P450 biocatalyst from the mycinamicin biosynthetic pathway. ACS Chem. Biol. 2016; 11:2642–2654. [PubMed: 27420774]
10. Li SY, Tietz DR, Rutaganira FU, Kells PM, Anzai Y, Kato F, Pochapsky TC, Sherman DH, Podust LM. Substrate recognition by the multifunctional cytochrome P450 MycG in mycinamicin

hydroxylation and epoxidation reactions. J. Biol. Chem. 2012; 287:37880–3780. [PubMed: 22952225]

11. Podust LM, Sherman DH. Diversity of P450 enzymes in the biosynthesis of natural products. Nat. Prod. Rpts. 2012; 29:1251–1266.
12. Poulos TL, Finzel BC, Howard AJ. High-resolution crystal structure of cytochrome P450cam. J. Mol. Biol. 1987; 195:687–700. [PubMed: 3656428]
13. OuYang B, Pochapsky SS, Dang M, Pochapsky TC. A functional proline switch in cytochrome P450(cam). Structure. 2008; 16:916–923. [PubMed: 18513977]
14. Asciutto EK, Dang M, Pochapsky SS, Madura JD, Pochapsky TC. Experimentally restrained molecular dynamics simulations for characterizing the open states of cytochrome P450(cam). Biochemistry. 2011; 50:1664–1671. [PubMed: 21265500]
15. Asciutto EK, Young MJ, Madura JD, Pochapsky SS, Pochapsky TC. Solution structural ensembles of substrate-free cytochrome P450cam. Biochemistry. 2012; 51:3383–3393. [PubMed: 22468842]
16. Colthart AM, Tietz DR, Ni Y, Friedman JL, Dang M, Pochapsky TC. Detection of substrate-dependent conformational changes in the P450 fold by nuclear magnetic resonance. Sci. Rep. 2016; 6:22035. [PubMed: 26911901]
17. Basudhar D, Madrona Y, Kandel S, Lampe JN, Nishida CR, de Montellano PRO. Analysis of Cytochrome P450 CYP119 Ligand-dependent conformational dynamics by two-dimensional NMR and X-ray crystallography. J. Biol. Chem. 2015; 290:10000–10017. [PubMed: 25670859]
18. Weigelt J. Single scan. sensitivity- and gradient-enhanced TROSY for multidimensional NMR experiments J. Am. Chem. Soc. 1998; 120:12706–12706.
19. Rückert M, Otting G. Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. J. Am. Chem. Soc. 2000; 122:7793–7797.
20. Giammona, D., Case, D., Bayly, C. Force field modifications for all-atom heme. 2001. Retrieved from http://research.bmh.manchester.ac.uk/bryce/amber
21. Freindorf M, Shao Y, Kong J, Furlani TR. Combined QM/MM calculations of active-site vibrations in binding process of P450cam to putidaredoxin. J. Inorg. Biochem. 2008; 102:427–432. [PubMed: 18180042]
22. Asciutto EK, Madura JD, Pochapsky SS, OuYang B, Pochapsky TC. Structural and dynamic implications of an effector-induced backbone amide cis-trans isomerization in cytochrome P450cam. J. Mol. Biol. 2009; 388:801–814. [PubMed: 19327368]
23. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. J. Am. Chem. Soc. 1982; 104:4546–4559.
24. Clore GM, Driscoll PC, Wingfield PT, Gronenborn AM. Analysis of the backbone dynamics of interleukin-1-beta using 2-dimensional inverse detected heteronuclear N-15-H-1 NMR spectroscopy. Biochemistry. 1990; 29:7387–7401. [PubMed: 2223770]
25. Voss NR, Gerstein M. 3V: cavity, channel and cleft volume calculator and extractor. Nucleic Acids Res. 2010; 38:W555–W562. [PubMed: 20478824]
26. Sligar SG. Coupling of spin, substrate and redox equilibria in cytochrome P450. Biochemistry. 1976; 15:5399–5406. [PubMed: 187215]
27. Wang A, Stout CD, Zhang QH, Johnson EF. Contributions of ionic interactions and protein dynamics to cytochrome P450 2D6 (CYP2D6) substrate and inhibitor binding. J. Biol. Chem. 2015; 290:5092–5104. [PubMed: 25555909]
28. Wester MR, Johnson EF, Marques-Soares C, Dansette PM, Mansuy D, Stout CD. Structure of a substrate complex of mammalian cytochrome P450 2C5 at 2.3 Å resolution: evidence for multiple substrate binding modes. Biochemistry. 2003; 42:6370–6379. [PubMed: 12767218]
29. Oshima R, Fushinobu S, Su F, Zhang L, Takaya N, Shoun H. Structural evidence for direct hydride transfer from NADH to cytochrome P450nor. J. Mol. Biol. 2004; 342:207–217. [PubMed: 15313618]
30. Monk BC, Tomasiak TM, Keniya MV, Huschmann FU, Tyndall JDA, O'Connell JD, Cannon RD, McDonald JG, Rodriguez A, Finer-Moore JS, Stroud RM. Architecture of a single membrane spanning cytochrome P450 suggests constraints that orient the catalytic domain relative to a bilayer. Proc. Natl. Acad. Sci. USA. 2014; 111:3865–3870. [PubMed: 24613931]

31. Denisov IG, Mak PJ, Grinkova YV, Bastien D, Berube G, Sligar SG, Kincaid JR. The use of isomeric testosterone dimers to explore allosteric effects in substrate binding to cytochrome P450 CYP3A4. J. Inorg. Biochem. 2016; 158:77–85. [PubMed: 26774838]
32. Denisov IG, Grinkova YV, Baylon JL, Tajkhorshid E, Sligar SG. Mechanism of drug-drug interactions mediated by human cytochrome P450 CYP3A4 monomer. Biochemistry. 2015; 54:2227–2239. [PubMed: 25777547]
33. Yao H, McCullough CR, Costache AD, Pullela PK, Sem DS. Structural evidence for a functionally relevant second camphor binding site in P450cam: Model for substrate entry into a P450 active site. Proteins: Struct., Funct., Bioinf. 2007; 69:125–138.
34. El-Sherbeni AA, El-Kadi AOS. Characterization of arachidonic acid metabolism by rat cytochrome P450 enzymes: The involvement of CYP1As. Drug Metab. Dispos. 2014; 42:1498–1507. [PubMed: 24969701]
35. Muller CS, Knehans T, Davydov DR, Bounds PL, von Mandach U, Halpert JR, Caflisch A, Koppenol WH. Concurrent cooperativity and substrate inhibition in the epoxidation of carbamazepine by cytochrome P450 3A4 active site mutants inspired by molecular dynamics simulations. Biochemistry. 2015; 54:711–721. [PubMed: 25545162]
36. Hartman JH, Knott K, Miller GP. CYP2E1 hydroxylation of aniline involves negative cooperativity. Biochem. Pharmacol. 2014; 87:523–533. [PubMed: 24345333]
37. Li S, Podust LM, Sherman DH. Engineering and analysis of a self-sufficient biosynthetic cytochrome P450 PikC fused to the RhFRED reductase domain. J. Am. Chem. Soc. 2007; 129:12940–12941. [PubMed: 17915876]
38. Lambalot RH, Cane DE, Aparicio JJ, Katz L. Overproduction and characterization of the erythromycin C-12 hydroxylase, EryK. Biochemistry. 1995; 34:1858–1866. [PubMed: 7849045]
39. Mendes MV, Anton N, Martin JF, Aparicio JF. Characterization of the polyene macrolide P450 epoxidase from Streptomyces natalensis that converts de-epoxypimaricin into pimaricin. Biochem. J. 2005; 386:57. [PubMed: 15228385]

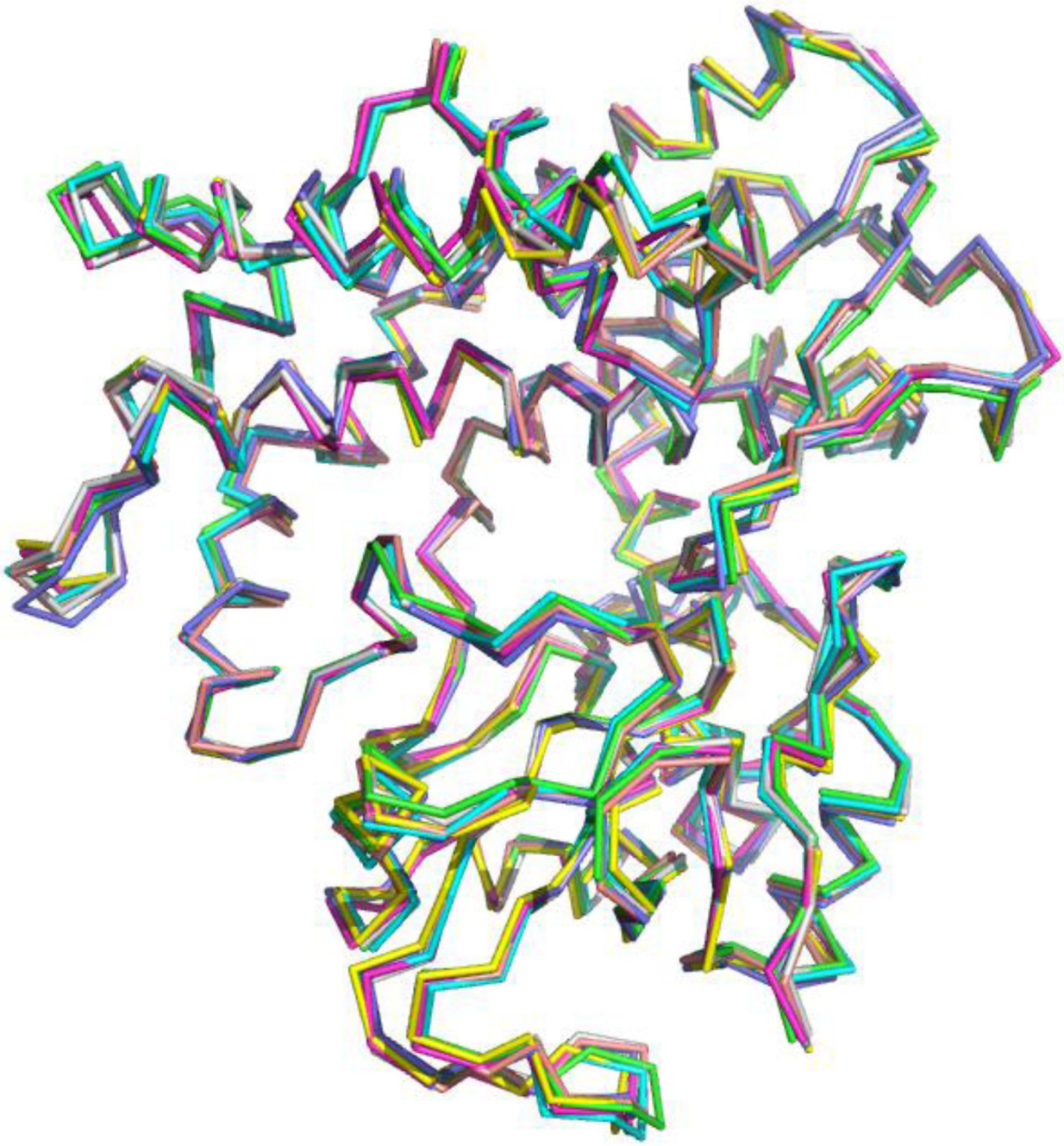


Figure 1.

Backbone alignment of seven M-IV-bound MycG coordinate sets showing the lowest energy penalties for RDC restraint violations. RMS deviation of backbone atoms is 0.618 Å. Molecular orientation is approximately the same as shown in Figure 2.

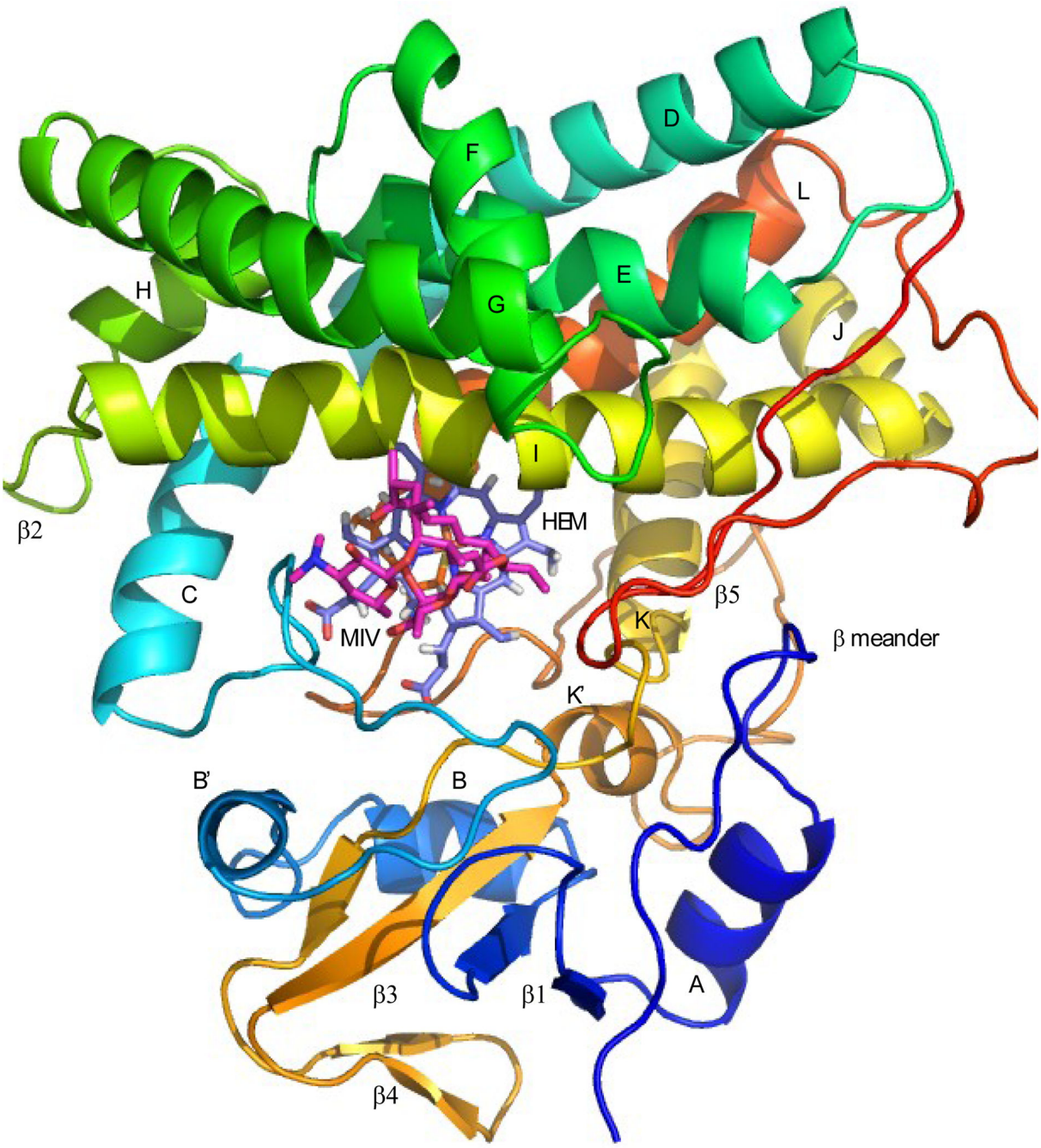


Figure 2.

Solution structure of M-IV-bound MycG (structure REP1), with secondary structural features labeled according to the convention of Poulos (12).Color progresses from blue (N-terminal) to red (C-terminal). Sequence numbering corresponds to that of entry 2Y98 in the RCSB PDB. Helix A, Gly 21-Thr 29; ᶫ1 (strand 1), Val 32-Val 35. (strand 2), Ala 43-Val 46; helix B, Tyr 49-Gly 57; helix B', Arg 63-Met 67; helix C, Pro 89-Ala 101; helix D, Ala 104-Thr 128; helix E, Asp 133-Leu 151; helix F, His 158-Ala 167; helix G, Ala 176-Lys 199; helix H, Leu 205-Gln 211; β2 (strand 1) Arg 213-Asp 214, (strand 2) Asp 217-Ser 218;

helix I, Glu 221-Thr 252; J helix, Pro 254-Asp 262; K helix, Pro 264-Trp 277; β3 (strand 1), Arg 288-Ala 290; β4 (strand 1), Val 294-Leu 296, (strand 2), Val 299-Ile 301; β3 (strand 2), Gly 304-Ala 309; helix K', Thr 311-Asn 315; helix L, Ala 349-Arg 366; β5 (strand 1), Leu 379-Gly 384, (strand 2), Arg 388-Pro 390. Heme is shown in light purple, MIV in crimson.

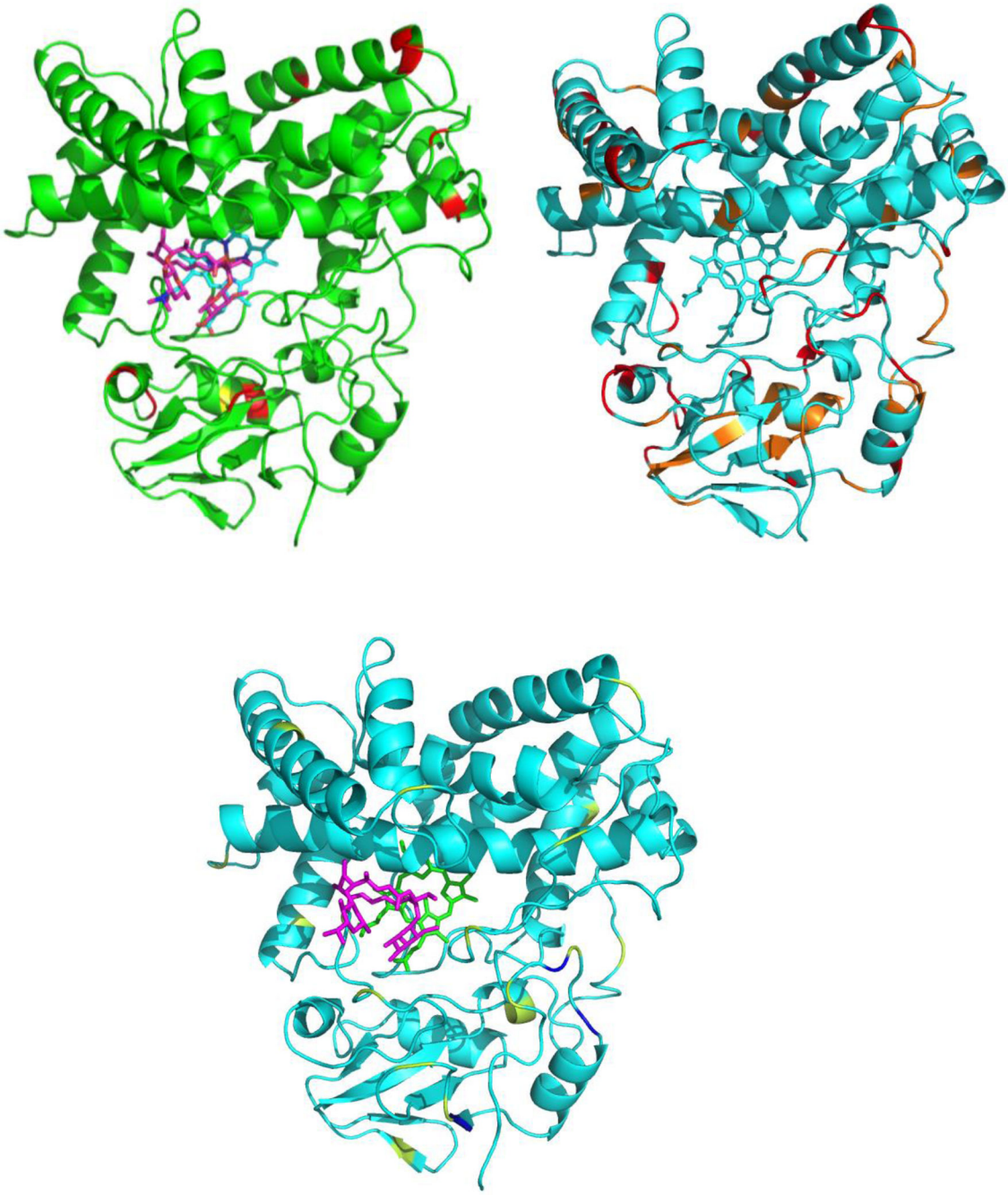


Figure 3.

NMR detected dynamic perturbations in MycG upon binding of M-IV. All structures are shown from the same orientation as in Figure 1. **Top right**: Spatial distribution of MygG backbone amide resonances most perturbed upon M-IV binding. Red, slow exchange behavior (NH correlations observed only at endpoints), **orange**, intermediate exchange (NH correlation observed shifted at each titration point). A complete tabulation of M-IV titration results is available in Supplementary Information. **Top left:** Spatial distribution of measurable hydrogen-deuterium exchange rates that differ significantly between M-IV-

bound and substrate-free MycG. Residues in red (Asp 52, Phe 61, Val 62, Ser 66, Met 67, Arg 73, Arg 115, Ala 126, Thr 128, Thr 252 and Val 395) show faster exchange in MIV-bound MycG, while Ala 54 (yellow) was the only identified residue that showed slower exchange in the presence of M-IV. **Bottom:** Spatial distribution of residues showing ^{15}N T_1/T_2 ratios more than one standard deviation above (blue) or below (yellow) the mean. See Table 2 for residue-specific details.

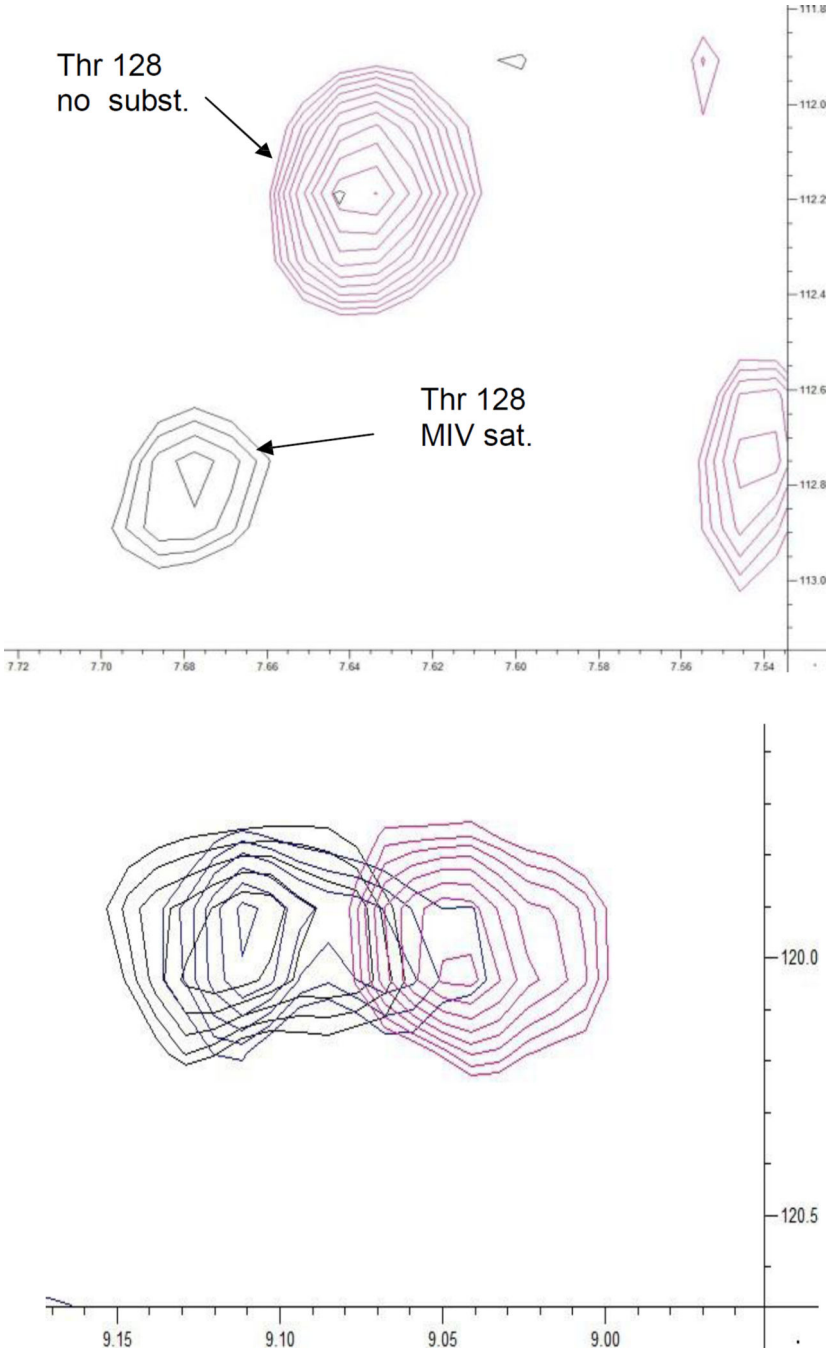


Figure 4.

(Top) N-H correlation of Thr 128 (C-terminal residue of helix D), which exhibits slow exchange behavior in the M-IV-MycG titration, monitored by $^{1}H,^{15}N$ TROSY-HSQC. The correlation is only observed in the absence of M-IV (red) and upon M-IV saturation (black). (Bottom) N-H correlation of Leu 84 (B'-C loop), showing intermediate exchange behavior. Substrate free spectrum is shown in red, 0.5/1 M-IV/MycG in gold and M-IV saturated MycG in black.

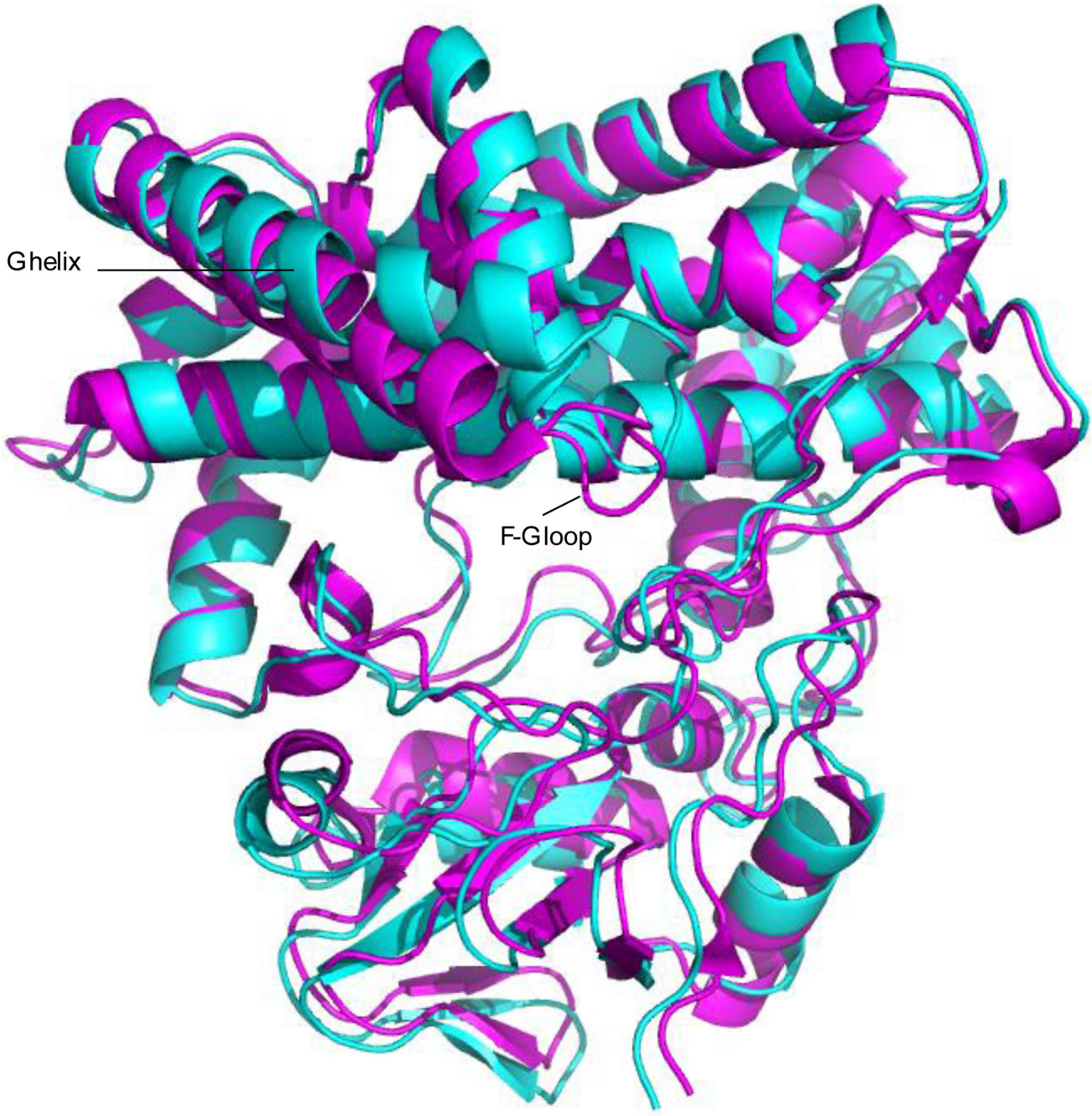


Figure 5.

Backbone superposition of 2Y98 [purple) and REP1 (cyan) structures of M-IV-bound MycG. The RMS deviation of the backbone atom positions of the structures is 1.49 Å. The largest differences are seen in the F-G loop and G helix (marked) with a concerted displacement of the β-rich region (towards the bottom of the structures as shown). Structures are shown in the same orientation as in Figure 2.

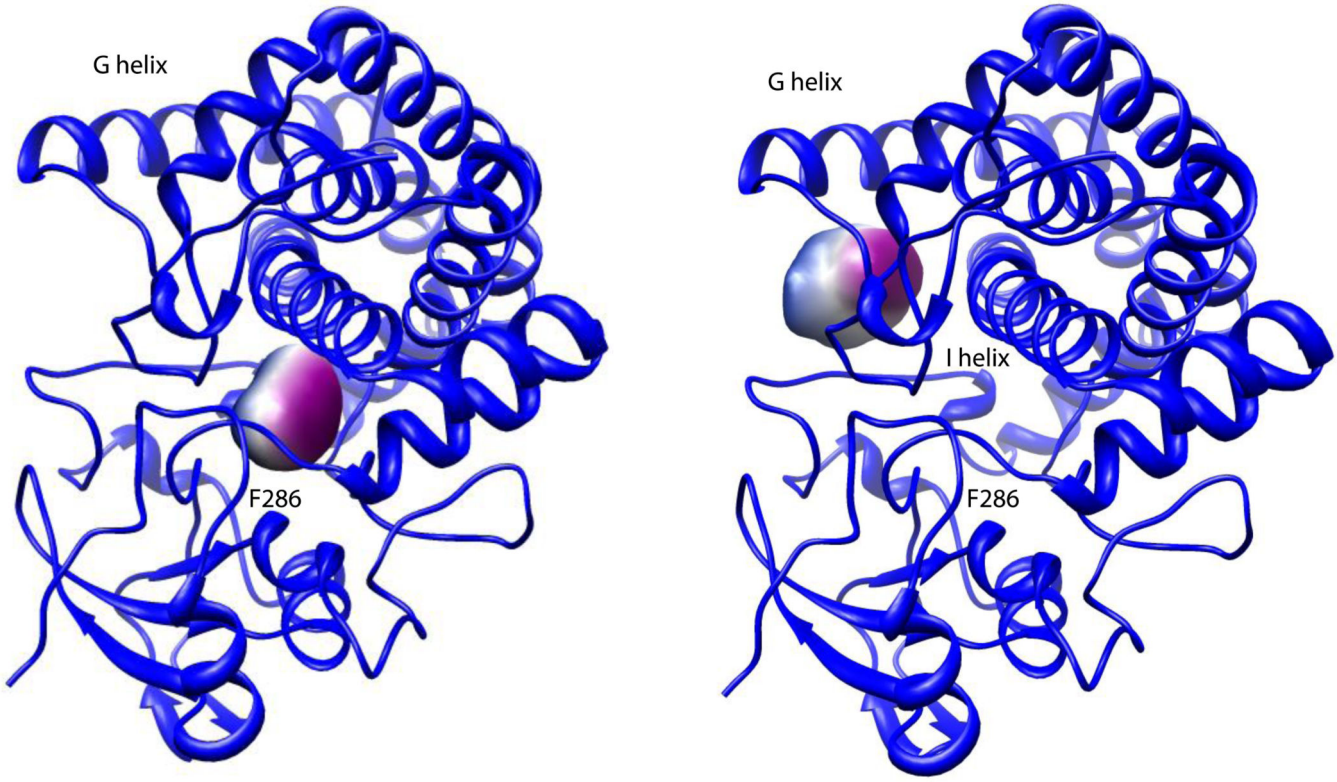


Figure 6.

Active site voids upon vacating M-IV from 2Y98. Two voids are detected, with the larger void (right) adjacent to the G helix and the F-G loop. The smaller void (left) is between the I helix and primary substrate contact Phe 286, immediately preceding the start of the β3 strand. Both figures are from the same perspective. Figures generated by V3 program (3vee.molmovdb.org). Void coloring is for visualization only. See text for details.

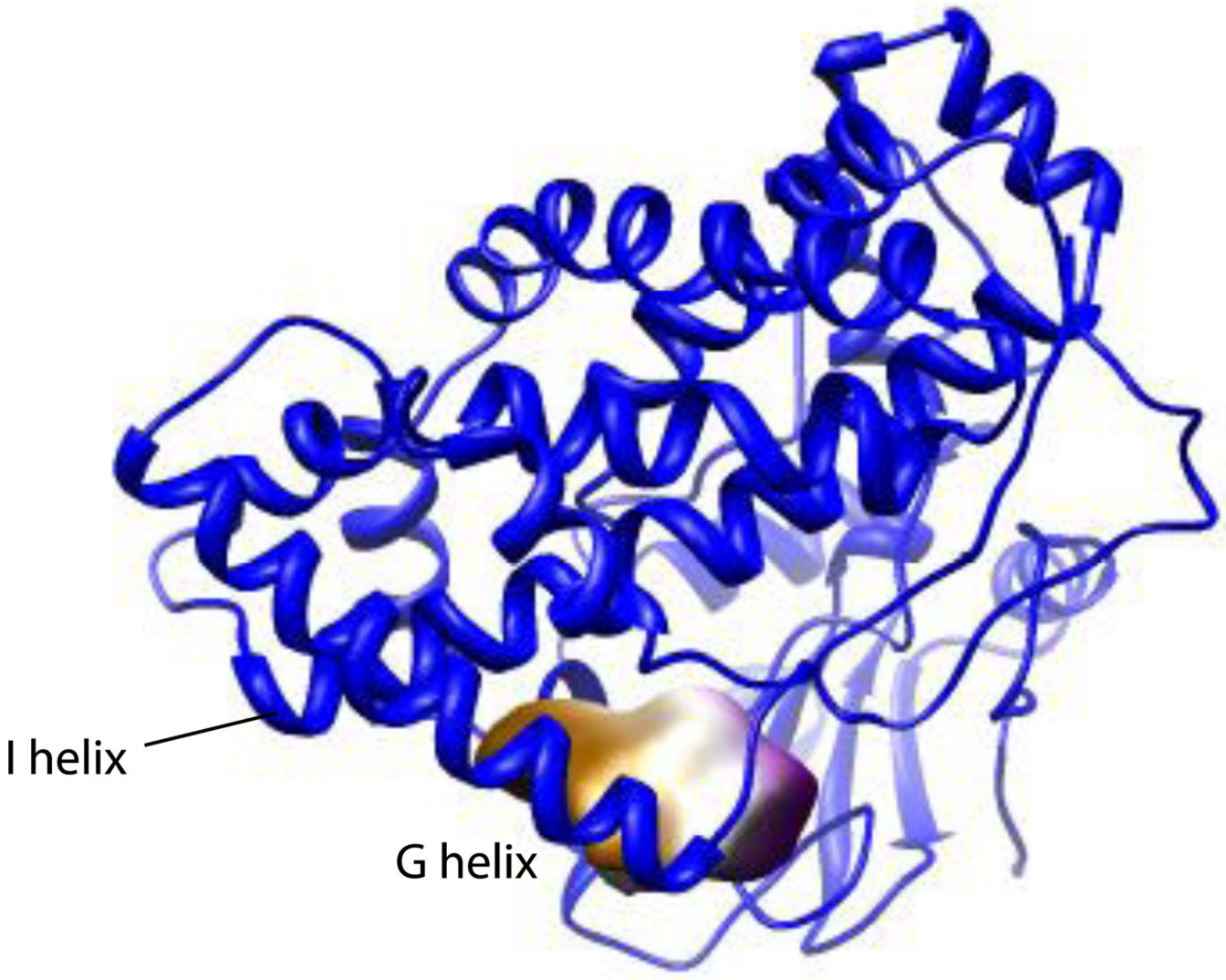


Figure 7.
Active site void upon vacating M-IV from REP1. Figure generated by V3 program (3vee.molmovdb.org). Void coloring is for visualization only. See text for details.

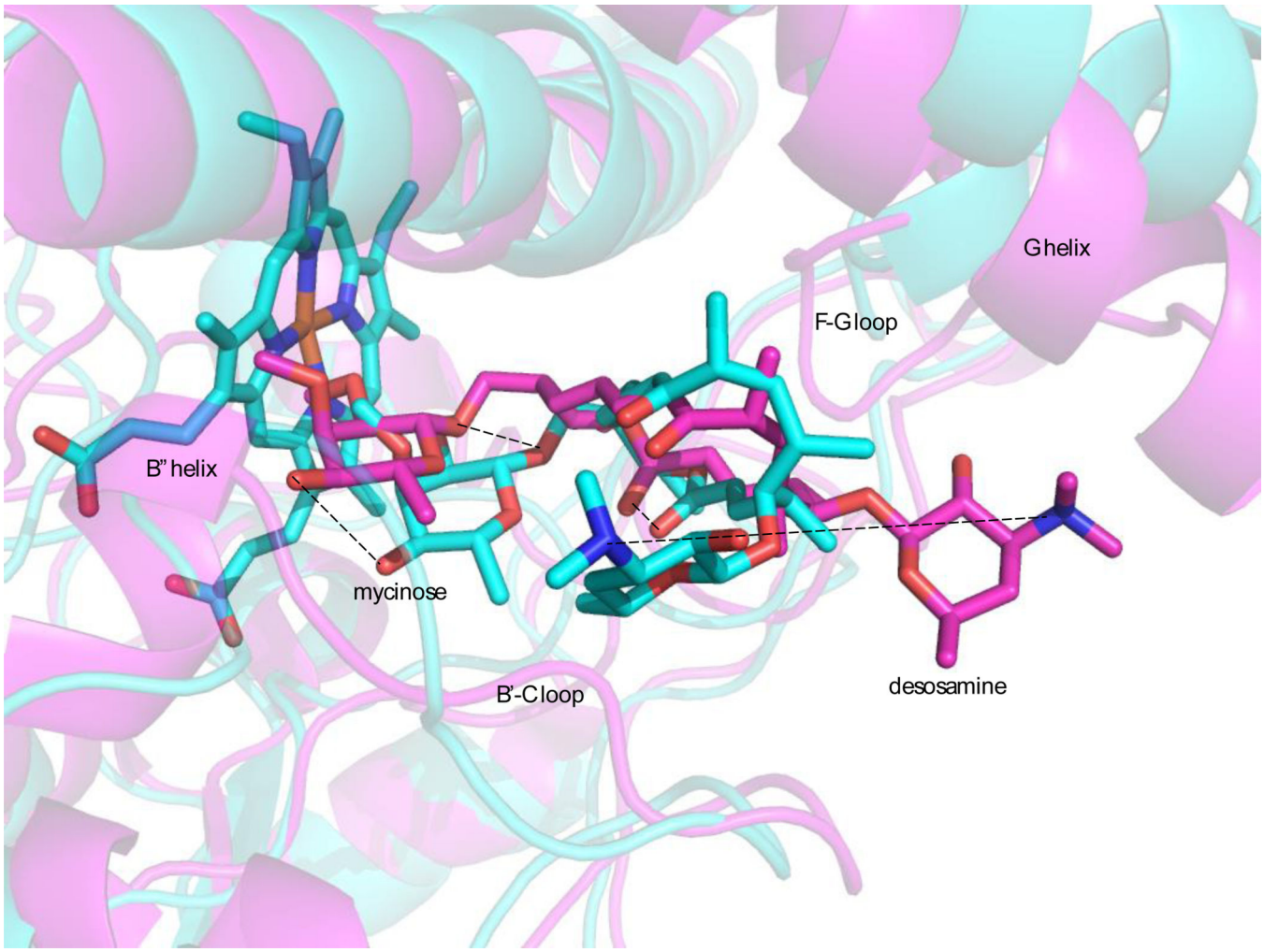


Figure 8.

Comparison of M-IV orientations in the active sites of REP1 (cyan) and 2Y98 (purple). M-IV adopts a “mycinose-in desosamine out” orientation in 2Y98, with the long axis of M-IV through the macrolactone ring roughly perpendicular to the heme plane and the desosamine close to the G helix and F-G loop. In REP1, the mycinose moves deeper into the active site pocket, placing the long axis of M-IV through the macrolactone ring roughly parallel to the heme plane and I helix and the desosamine ring close to the B’-C loop. Dotted lines connect corresponding M-IV atoms in both structures.

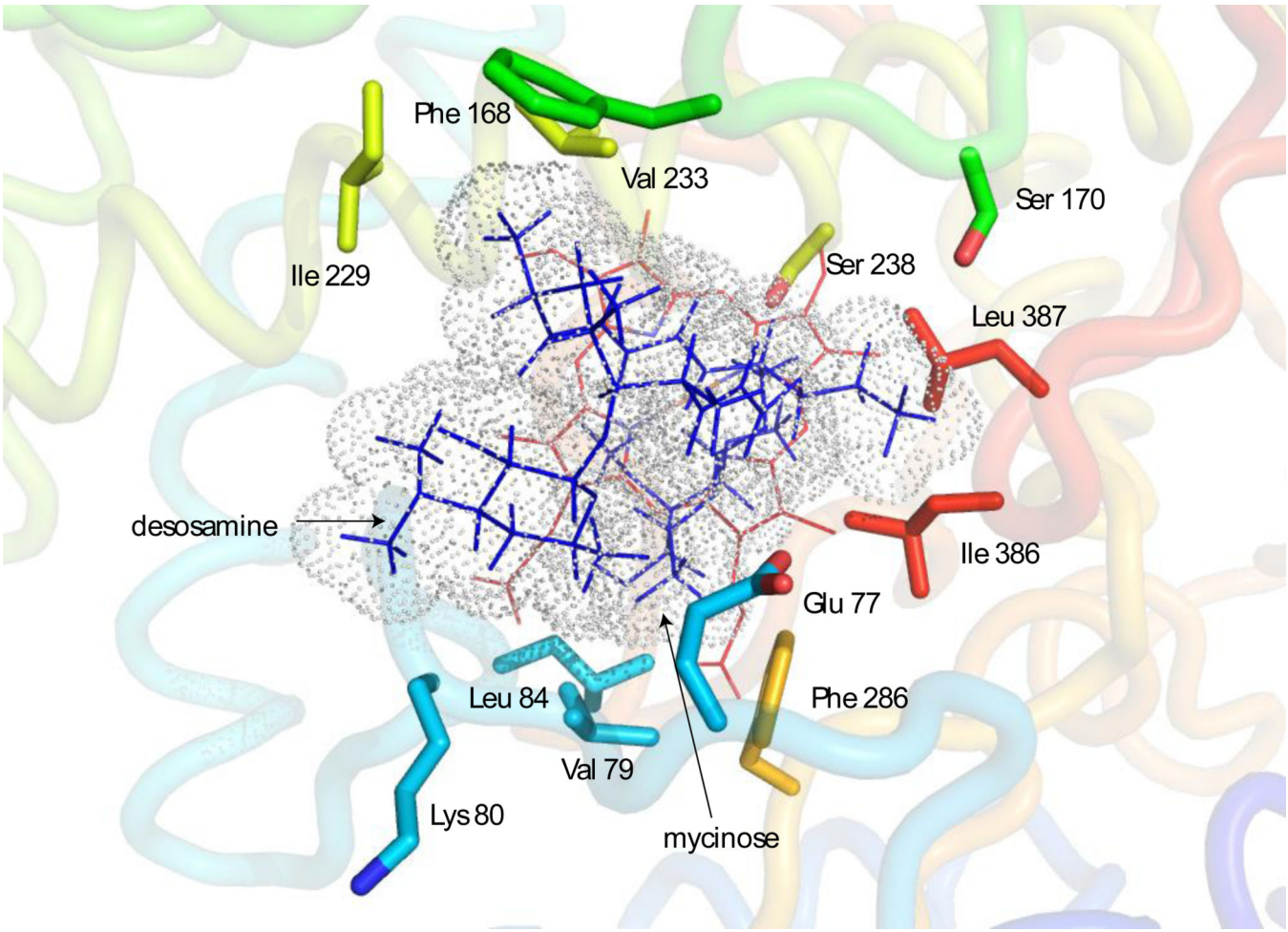


Figure 9.

Side chains of active site residues making contact with M-IV in the REP1 structure. M-IV is represented with blue lines and van der Waals radii as dots. Heme is shown in red lines. Only heavy atoms of contact side chains are represented. See text for details.

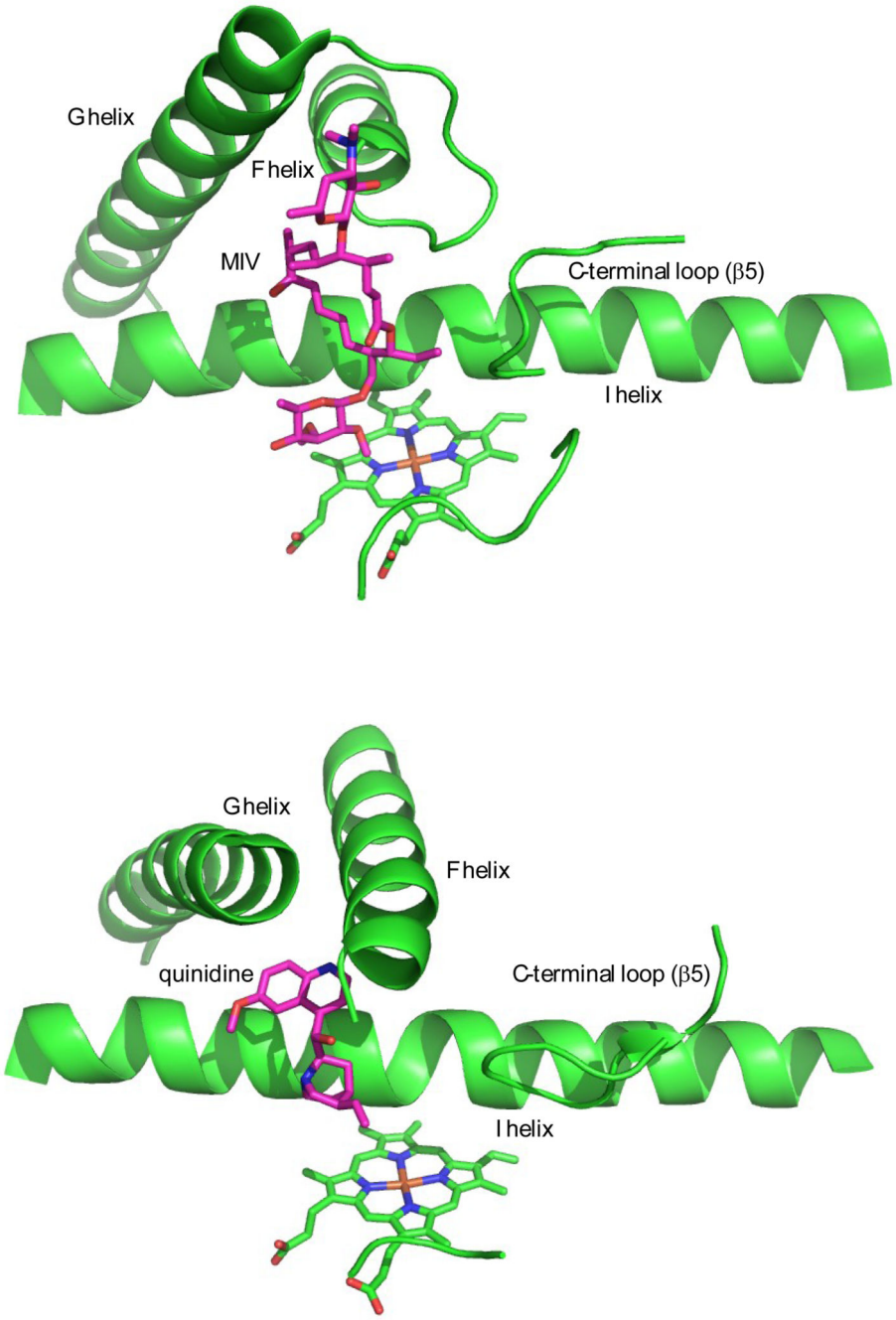
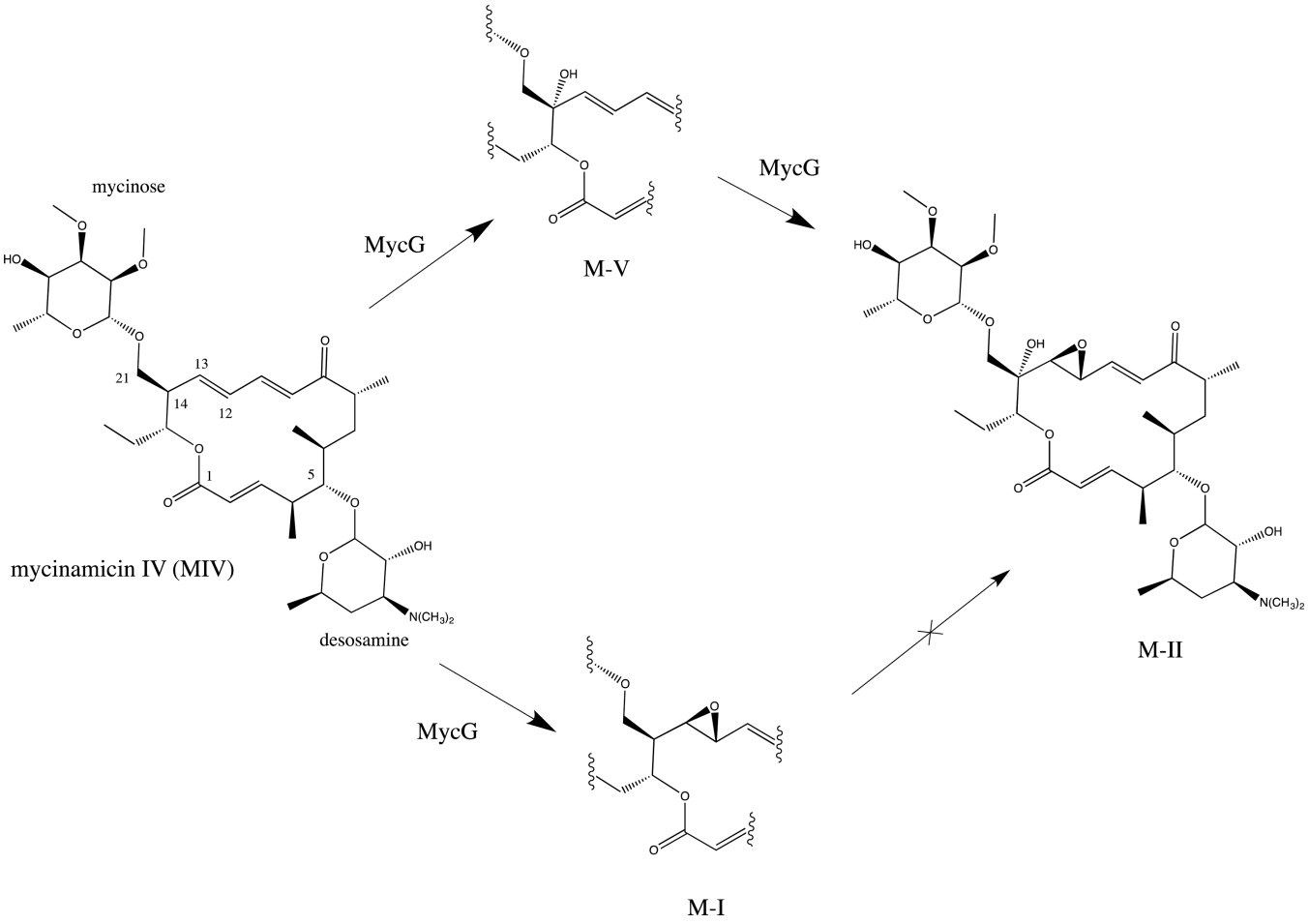


Figure 10.

Comparison of crystallographic M-IV binding orientation in MycG (top, PDB entry 2Y98) with that of inhibitor quinidine in CYP2D6 (bottom, PDB entry 4WNU, ref. (27)). Molecules are oriented so that the heme plane and orientation is in approximately the same tilt and direction in both structures.



Scheme 1.

Oxidations of M-IV catalyzed by MycG. Numbering of M-IV skeleton is as referred to in the text.

Table 1

Concentrations of MycG and M-IV for each point of the titration of MycG with M-IV.

[MycG] μM	[M-IV] μM	Ratio M-IV/MycG
460	92	0.2
460	184	0.4
460	276	0.6
460	368	0.8
460	460	1.0
460	920	2.0

Table 2

M-IV-bound MycG residues with T_1/T_2 ratios more than one standard deviation (SD) from the mean, with secondary structural location noted.

T_1/T_2 > 1 SD above mean (ms contribution)	**T_1/T_2 > 1 SD below mean (ps-ns contribution)**
Gly 16 (N-terminal loop)	Ser 3 (N-terminal loop)
Ala 20 (A helix)	Ala 4 (NT-terminal loop)
Arg 34 (β1, strand 1)	Val 14 (N-terminal loop)
Val 35 (loop, β1)	Leu 17 (N-terminal loop)
Glu 41 (loop,β1)	Arg 36 (loop,β1)
Glu 42 (loop,β1)	Arg 39 (loop,β1)
Asp 51 (beginning of B helix)	Glu 77 (B'-C loop)
Gly 59 (B-B' loop)	Leu 83 (B'-C loop)
Val 62 (B-B' loop)	Ser 92 (C helix)
Arg 63 (B-B' loop)	Gly 129 (D-E loop)
Leu 84 (B' loop)	Val 174 (E-F loop)
Asp 123 (D helix)	Tyr 189 (F helix)
Val 307 (β3 adjacent to β1)	Asp 217 (H-I loop)
Asp 328 (β meander)	Leu 266 (J-K loop)
Gly 369 (L-β5 loop)	Thr 300 (β4)
Gly 373 (L-β5 loop)	Ala 313 (K' helix)
Val 395 (C-terminal loop)	Ala 314 (K' helix)
Trp 396 (C-terminal residue)	Arg 326 (β meander)
	Gly 384 (β5 turn)
	Glu 392 (C-terminal loop)