

Conformational Analysis of Free and Bound Retinoic Acid

Zheng Fu, Xue Li, and Kenneth M. Merz, Jr.*

Department of Chemistry and the Quantum Theory Project, 2328 New Physics Building, P.O. Box 118435, University of Florida, Gainesville, Florida 32611-8435, United States

Supporting Information

ABSTRACT: The conformational profiles of unbound all-*trans* and 9-*cis* retinoic acid (RA) have been determined using classical and quantum mechanical calculations. Sixty-six all-*trans*-RA (ATRA) and 48 9-*cis*-RA energy minimum conformers were identified via HF/6-31G* geometry optimizations *in vacuo*. Their relative conformational energies were estimated utilizing the M06, M06-2x, and MP2 methods combined with the 6-311+G(d,p), aug-cc-pVDZ, and aug-cc-pVTZ basis sets, as well as complete basis set MP2 extrapolations using the latter two basis sets. Single-point energy calculations performed with the M06-2x density functional were found to yield similar results to MP2/CBS for the low-energy retinoic acid conformations. Not unexpectedly, the conformational propensities of retinoic acid were governed by the orientation and arrangement of the torsion angles associated with the polyene tail. We also used previously reported QM/MM X-ray refinement results on four ATRA-protein crystal structures plus one newly refined 9-*cis*-RA complex (PDB ID: 1XDK) in order to investigate the conformational preferences of bound retinoic acid. In the rerefined RA conformers, the conjugated double bonds are nearly coplanar, which is consistent with the global minimum identified by the Omega/QM method rather than the corresponding crystallographically determined conformations given in the PDB. Consequently, a 91.3% average reduction of the local strain energy in the gas phase, as well as 92.1% in PCM solvent, was observed using the QM/MM refined structures versus the PDB deposited RA conformations. These results thus demonstrate that our QM/MM X-ray refinement approach can significantly enhance the quality of X-ray crystal structures refined by conventional refinement protocols, thereby providing reliable drug-target structural information for use in structure-based drug discovery applications.

INTRODUCTION

All-*trans* retinoic acid (ATRA) and 9-*cis* retinoic acid (9-*cis*-RA) are derivatives of vitamin A. They are involved in many crucial biological processes by modulating the communication between their target retinoic acid receptors (RARs) and retinoid X receptors (RXRs), in the intracellular environment. In the ligand-free (apo) state, RAR and RXR are prone to form a RAR–RXR heterodimer complex that allows dual ligand input. When ATRA binds to RAR as an agonist, it alters the RAR conformation *via* an allosteric effect and removes the corepressor (CoR) bound on the RAR surface. Then, in the presence of 9-*cis*-RA a coactivator (CoA) interaction area is exposed at the RXR subunit that facilitates the recruitment of CoA to the RAR–RXR heterodimer cooperatively. As a result, one or more transcription/epigenetic machineries are activated and bind to the target gene promoter regions, and the RAR–RXR heterodimer itself is also able to bind to cognate *cis*-acting DNA regulatory elements (termed RA response elements) *via* its DNA binding domain.¹

As members of the nuclear receptor (NR) family, RAR and RXR mediate many signaling pathways that control cellular proliferation, differentiation, and the early stages of carcinogenesis. Hence, both ATRA and 9-*cis*-RA have served as drug molecules in cancer therapy and prevention, predating the determination of the RAR and RXR crystal structures.² Clinical trials have confirmed that ATRA can induce complete remission and overall survival in most acute promyelocytic leukemia (APL) patients.^{3–5} In APL cells, RAR α is fused with the promyelocytic leukemia (PML) gene and produces the RAR α –PML fusion protein.⁶ This hybrid protein not only tightly binds and recruits the nucleosome remodeling and deacetylase (NuRD) enzyme

that facilitates polycomb repressive complex 2 (PRC2) binding to block retinoic acid signaling⁷ but changes both the structures and intracellular localizations of RXR and other nuclear antigens as well.⁶ The chaotic and abnormal nuclear organization in APL cells ultimately leads to a few specific chromosomal translocations and triggers acute promyelocytic leukemia by arresting myelopoiesis at the promyelocyte stage. Nevertheless, *in vitro* RA treatment can reorganize the subnuclear localization of RAR α -PML with RXR and other essential nuclear proteins which in turn eliminates the chromosomal translocations.⁶ Meanwhile therapeutic doses of ATRA restore the normal expression level of tumor-suppressor genes for APL patients *via* transcriptional corepressor dissociation from the promoter region of genes like RAR β_2 .⁷ In addition, numerous animal experiments demonstrate that ATRA and 9-*cis*-RA are efficacious in abrogating tumor progression in models of breast,^{8–10} hepatic,¹¹ and pancreatic¹² carcinogenesis.

Although both ATRA and 9-*cis*-RA are potent anticancer agents, their side effects limit the widespread use of these drugs in human cancer treatment, which includes teratogenicity,^{13,14} severe headaches, and hepatotoxicity.¹⁵ Therefore, a comprehensive conformational analysis of retinoic acid will enable the rational design of more selective retinoids (ATRA and its analogs) and rexinoids (9-*cis*-RA and its analogs) in order to attenuate adverse side effects. To the best of our knowledge, this analysis has not been performed to date at a high level of sophistication. Beppu and Katitani used the CNDO/2 method to explore the adiabatic potential surface (APS) of ATRA.¹⁶

Received: November 12, 2011

Published: February 24, 2012



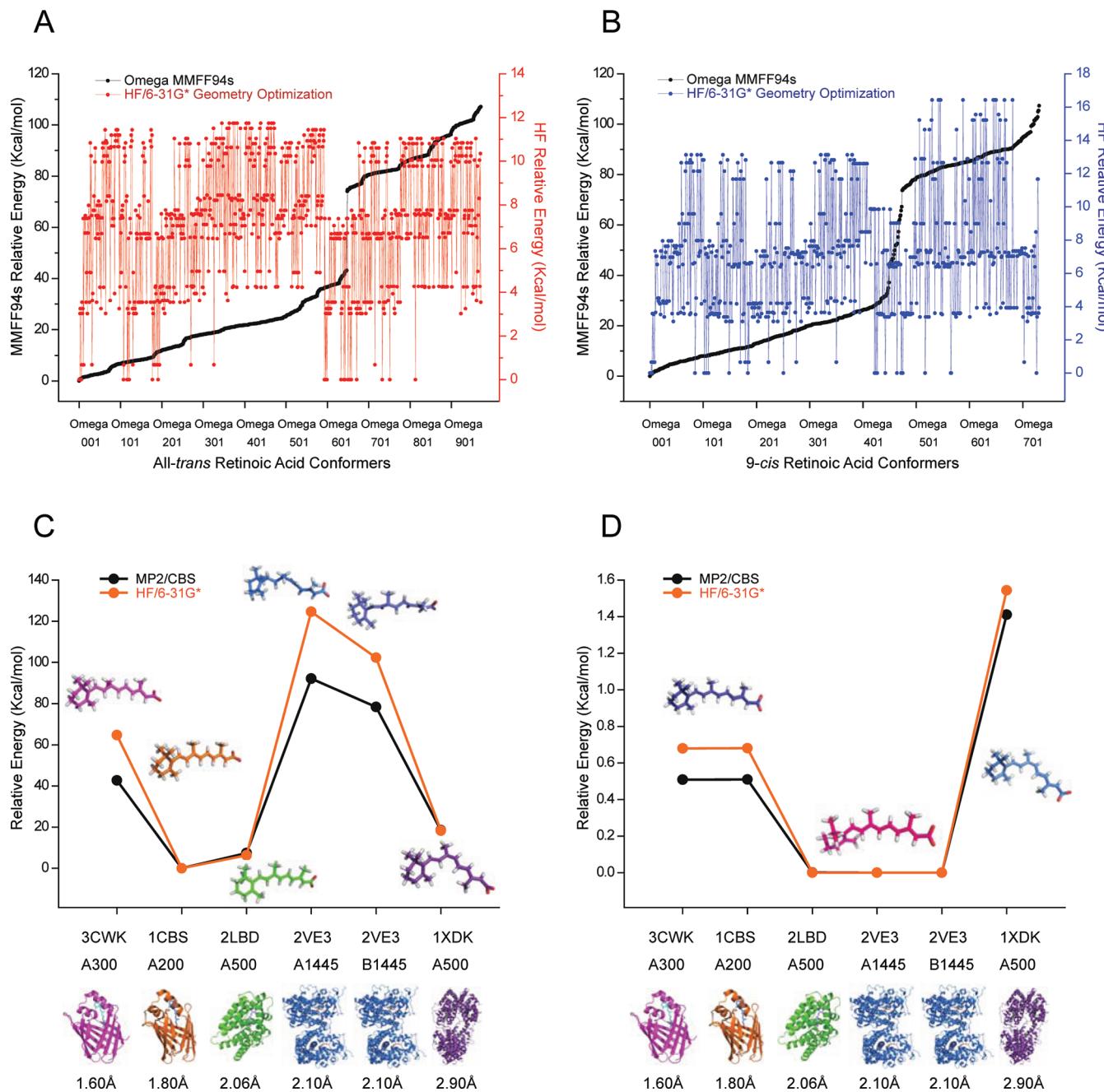


Figure 1. Omega generated retinoic acid conformational ensembles optimized by HF/6-31G* for (A) *all-trans* and (B) *9-cis* conformers. Relative energy of the retinoic acid conformers deposited in the PDB (C) and their optimized geometries using HF/6-31G* in the gas phase (D).

According to their results, the ATRA adiabatic potential surface has an energy maximum when Φ_1 (C5–C6–C7–C8) is around -170° with CNDO/2 geometry optimization (for the atom numbering for ATRA, see Figure 2A). Furthermore, they calculated the APS of several ATRA analogs to evaluate the conformational flexibility of the β -cyclohexenyl ring, with particular attention given to puckered forms. Six years later, Aalten and co-workers generated an exhaustive ensemble of ATRA conformers to calculate their relative conformational energies via MD simulation. They noted that ATRA bound to cRBP (cellular retinol-binding protein, type I) is highly strained relative to free ATRA conformers.¹⁷ Although these studies are helpful in elucidating the conformational space of retinoic acid, they neither considered *9-cis*-RA nor did they analyze the conformational profiles of retinoic acids bound to the RAR–

RXR heterodimer. Hence, we carried out a thorough conformational analysis of both ATRA and *9-cis*-RA in their free and bound states with highly accurate MP2/CBS calculations. QM/MM X-ray refinement was also performed on RA within the RAR/RXR active site. For more detail regarding the QM/MM refined RA–protein complexes, please refer to Li and co-workers.¹⁸

RESULTS AND DISCUSSIONS

Conformational Preferences of All-trans and 9-cis-RAs in the Gas Phase. Geometry optimizations at the HF/6-31G* level of theory were initiated from 948 OMEGA generated ATRA structures (described in the Materials and Methods section), which resulted in the identification of 66 unique low-energy conformers (Figure 1A). Similarly, we generated the *9-cis*-RA OMEGA geometries and optimized them at the HF/6-31G* level and

identified 48 unique low-energy structures (Figure 1B). To explore RA conformations in the context of their target binding sites, we selected six different retinoic acid conformers (five ATRA and one 9-*cis*-RA) from five PDB entries (see Materials and Methods section) and performed single point energy calculations, as well as geometry optimizations, at the HF/6-31G* level of theory in the gas phase. Note that along the *x* axis the RA–protein complexes are listed in ascending order of their resolution. MP2/CBS calculations were used as a “gold standard” to estimate the single point energies of the PDB deposited retinoic acid conformers and their respective optimized structures. Figure 1C shows that the maximum conformational energy interval (using the 1CBS A200 ATRA conformer as reference) is more than 120 kcal/mol, yet after *ab initio* QM optimization, only three low energy structures were identified (two ATRA and one 9-*cis*-RA, Figure 1D). This result reveals that there are significant coordinate errors in bound retinoic acid conformers, giving rise to large ligand strain energies.

All-*trans* retinoic acid can adopt different conformations primarily via sampling the torsion angles Φ_1 (C5–C6–C7–C8), Φ_2 (C7–C8–C9–C10), Φ_3 (C9–C10–C11–C12), Φ_4 (C11–C12–C13–C14), and Φ_5 (C13–C14–C15–O1) on its polyene tail (Figure 2A). We use the notation Φ (C_g – Ch – Ci – Cj) to stand for the torsion angle between the g – h – i and h – i – j planes. The positive sign means clockwise rotation of the h – i bond looking from atom h to atom i , whereas the minus sign indicates a counterclockwise rotation. Figure 2B shows that the 66 ATRA low energy conformers favor either Φ_1 (C5–C6–C7–C8) $\approx -63^\circ$ or $+63^\circ$. This torsion angle reflects the orientation of the ionone ring relative to the polyene chain, and its X-ray crystallography experimental value¹⁹ is $+59^\circ$. Unlike the C6–C7 bond, the rotation of the C8–C9, C10–C11, and C12–C13 bonds imparts more conformational flexibility on the polyene tail, where Φ_2 (C7–C8–C9–C10) $\approx \pm 41^\circ$ or $\pm 180^\circ$, Φ_3 (C9–C10–C11–C12) $\approx \pm 45^\circ$ or $\pm 180^\circ$, and Φ_4 (C11–C12–C13–C14) $\approx \pm 38^\circ$ or $\pm 180^\circ$. Figure 2B–E show that the values of Φ_1 , Φ_2 , Φ_3 , and Φ_4 are discretely distributed, yet in Figure 2F we observe substantial scattering from -180° to $+180^\circ$ with irregular relative energies. This indicates that the C14–C15 bond has more rotational freedom than its counterparts in the polyene tail, which was noted by Klucik and co-workers from a different perspective.²⁰ They developed a novel methodology (TACS) to search the ligand bound conformation and employed retinoic acid as a probe to evaluate its predictive power. Their result shows that the C13–C14–C15–O1 torsion makes the maximum contribution to the RMSD (around 16°) when compared with the putative bioactive conformations in RA–protein complexes.²⁰ Another interesting point is that the Φ_5 (C13–C14–C15–O1) torsion angle at the global minimum is not equal to 0° or 180° . The exact Φ_5 value at the ATRA global minimum is around 22° , and its neighbors on the potential energy surface (relative energy less than 0.7 kcal/mol, see group one in Figure 3) have Φ_5 values in a range between 29° and 33° . One possible explanation of this observation is the steric hindrance of the C20 methyl group. Moreover, carbon is more electronegative than hydrogen; hence, C20 might have a negative charge (Mulliken atomic charge = -0.828918 at the ATRA global minimum). Note that we use the anionic form of retinoic acid in all calculations, so there might be an electrostatic repulsion between the C20 atom and the carboxyl group, which causes the preferred C14–C15 orientation where the oxygen atom staggers the C20 methyl group. These two synergistic effects

yield a Φ_5 (C13–C14–C15–O1) torsion angle that is not equal to 0° or 180° .

If we only consider the absolute values of Φ_1 – Φ_4 (Φ_5 was excluded due to its extreme variation), the 66 ATRA low-energy conformers can be classified into 1 ($|\Phi_1| \approx 63^\circ$) \times 2 ($|\Phi_2| \approx 41^\circ$ or 180°) \times 2 ($|\Phi_3| \approx 45^\circ$ or 180°) \times 2 ($|\Phi_4| \approx 38^\circ$ or 180°) = 8 energetically distinct conformational groups. Figure 3 shows the typical backbone structure of each group with their corresponding average relative energy (reference is the Omega plus HF/6-31G* identified global minimum). The standard deviation within each group is presumably ascribed to the rotation of the C14–C15 bond (via the diversity of Φ_5 value). From Figure 3, we can also infer that the relative energy ordering of different ATRA conformer groups depends on the adoption of the torsion angles associated with the polyene tail. There are several cases to consider: First, when the four double bonds (C7=C8, C9=C10, C11=C12, and C13=C14) are coplanar, which indicates $|\Phi_2| \approx |\Phi_3| \approx |\Phi_4| \approx 180^\circ$, ATRA conformers will have the smallest relative energies and are near global minimum on the conformational energy surface (see group one in Figure 3). The C5=C6 double bond is not counted in this energy-structure relationship analysis since the $|\Phi_1|$ value is a constant ($\approx 63^\circ$). In the next instance, if three conjugated double bonds are in the same plane ($|\Phi_2| \approx |\Phi_3| \approx 180^\circ$ or $|\Phi_3| \approx |\Phi_4| \approx 180^\circ$, like groups two and three in Figure 3), the relative energies increase by 3–4 kcal/mol. Another set has two pairs of conjugated double bonds, and within each pair the double bonds are coplanar yet the entire set of four double bonds does not lie in the same plane ($|\Phi_2| \approx |\Phi_4| \approx 180^\circ$, $|\Phi_3|$ is not equal to 0° or 180° : see group four in Figure 3). In this scenario, the energy gap between the ATRA low-energy conformers and the most stable conformation of the set is around 4–5 kcal/mol. Next, if only two conjugated double bonds are coplanar ($|\Phi_2| \approx 180^\circ$ or $|\Phi_3| \approx 180^\circ$ or $|\Phi_4| \approx 180^\circ$, see groups five, six, and seven in Figure 3), we obtain ATRA structures that are about 7–8 kcal/mol higher in energy than the lowest energy conformer we identified. The final situation has none of the four conjugated double bonds in the same plane ($|\Phi_2| \neq |\Phi_3| \neq |\Phi_4| \neq 180^\circ$), and these ATRA conformers yield the highest relative energies (see group eight in Figure 3).

In addition to subdividing the 66 ATRA low-energy conformers in terms of their polyene chain torsion angles, 3D similarity comparisons by ROCS between each pair were performed to offer another criterion for categorizing these structures. The 4356 Tanimoto scores were clustered into four smaller sets by virtue of the centroid linkage algorithm with estimating the uncentered correlation as the similarity metric. This clustering result is visualized in Supporting Information Figure 1A. Even though our clustering method is not one-size-fits-all and there are biases on the clusters it constructs, we can still observe that most of the ATRA low-energy conformers in groups one, two, and three fall into cluster II, and those in group four into cluster I. Clusters III and IV mainly cover the structures in groups five, six, seven, and eight. This clustering result demonstrates that our conformational analysis based on Φ_1 – Φ_5 dihedral angles is reasonable in describing the conformational preference of free ATRA in the gas phase.

We also carried out a similar conformational analysis procedure on 9-*cis*-RA, and the results are presented in Figures 4 and 5. The dihedral angles Φ_1 (C5–C6–C7–C8) and Φ_4 (C11–C12–C13–C14) from 9-*cis*-RA low-energy conformers are approximately equal to those found in low energy ATRA structures ($\Phi_1 \approx \pm 63^\circ$, $\Phi_4 \approx \pm 38^\circ$ or $\pm 180^\circ$). Yet the values of

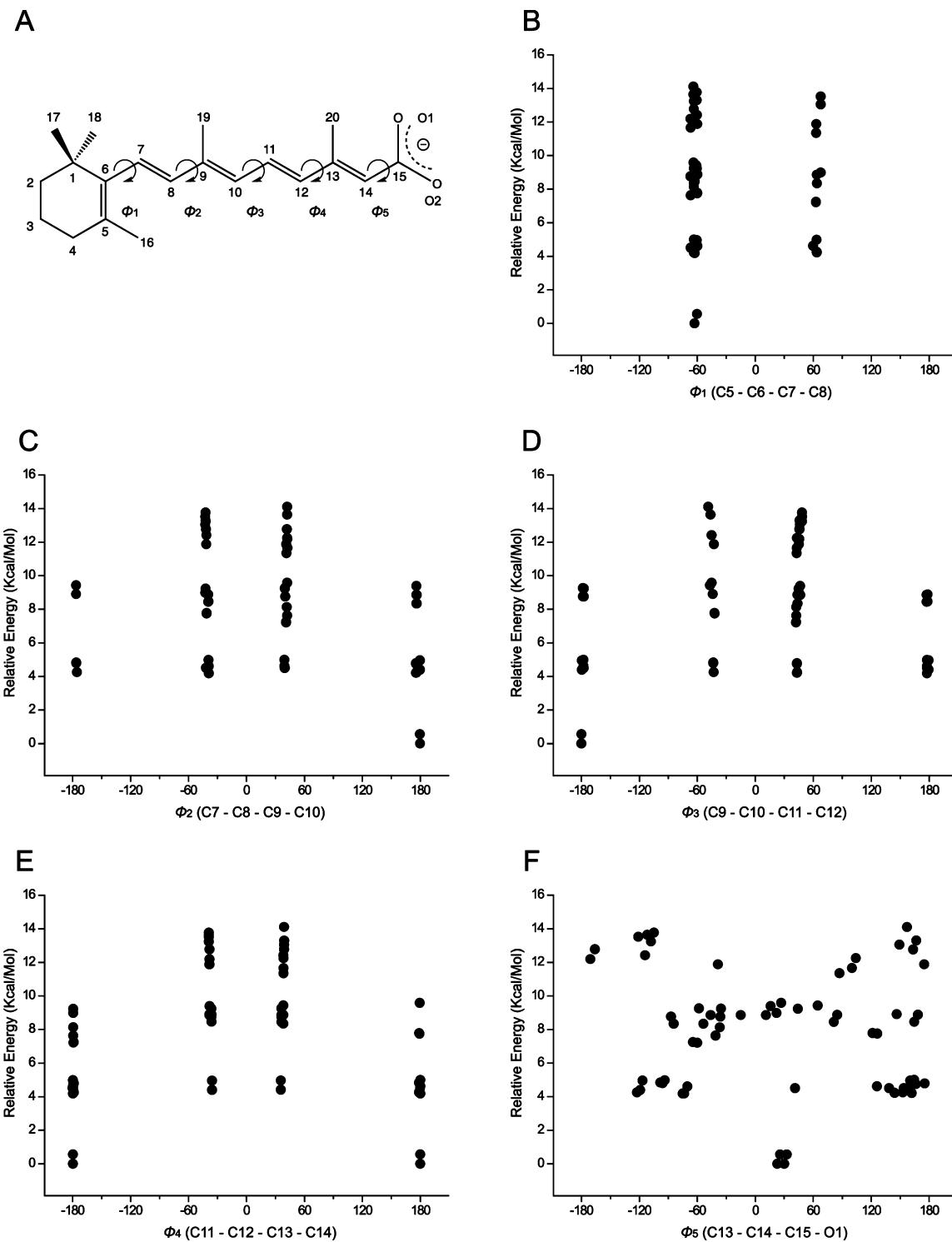


Figure 2. (A) Structure of all-trans retinoic acid (ATRA). The numbers of carbon and oxygen atoms are labeled. (B–F) Relative energies of 66 ATRA low-energy conformers as a function of different dihedral angles (in degrees) along the polyene chain. Relative energies were calculated at the MP2/aug-cc-pVDZ level in the gas phase.

Φ_2 ($C_7-C_8-C_9-C_{10}$) and Φ_3 ($C_9-C_{10}-C_{11}-C_{12}$) are slightly different from their counterparts in ATRA. For Φ_2 ($C_7-C_8-C_9-C_{10}$), it prefers either $\pm 48^\circ$ or $\pm 180^\circ$, and Φ_3 ($C_9-C_{10}-C_{11}-C_{12}$) adopts values of $\pm 50^\circ$ or $\pm 180^\circ$ (Figure 4C, D, and E). As expected, the rotational freedom of the $C_{14}-C_{15}\sigma$ bond is much greater than any other single bond in the polyene tail, and the global minimum Φ_5 value is close to but not equal to 0° (Figure 4F). Similarly, we can

classify the 48 9-cis-RA minimum energy conformations into eight groups in light of the absolute values of Φ_2 , Φ_3 , and Φ_4 . 9-cis-RA conformations with all conjugated double bonds (excluding $C_5=C_6$) in the same plane have the lowest relative energies and include the global minimum (group one in Figure 5). The less favorable orientation of the 9-cis-RA polyene tail has three coplanar double bonds and has an energy penalty of 3.5–3.8 kcal/mol (groups two and three in Figure 5). 9-cis-RA

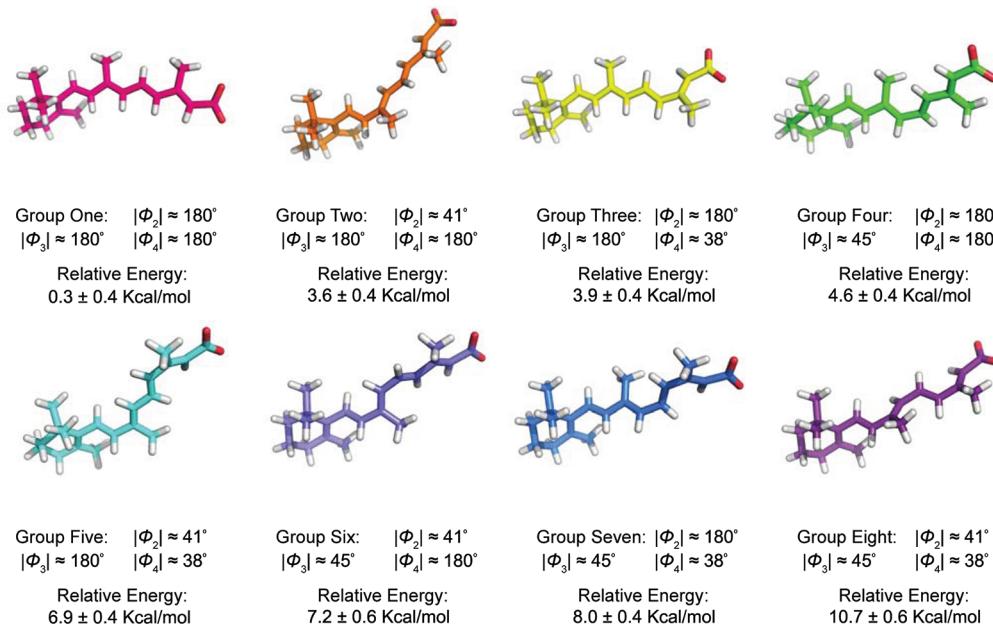


Figure 3. Schematic representation of the eight groups of ATRA low energy conformers identified by the Omega/*ab initio* QM method. Groups are arranged in the order of increasing average relative energy obtained via MP2/aug-cc-pVDZ calculations in the gas phase.

conformers that have Φ_3 at $\pm 50^\circ$ while keeping $|\Phi_2| \approx |\Phi_4| \approx 180^\circ$ are separated from the global minimum by nearly 4 kcal/mol (group four in Figure 5). Conformations in groups four to seven have relatively similar energy-structure relationships to those found for ATRA and are destabilized by 6.9 to 9.4 kcal/mol. Ultimately, geometries in group eight exhibit an average energy value more than 10 kcal/mol above the global minimum, which contributes to their highly distorted backbone structure.

Supporting Information Figure 1B displays the clustering of the Tanimoto scores among the 48 9-*cis*-RA minimum energy conformations. Cluster I includes all four members of group one, and cluster II includes the majority of group two conformers. Geometries belonging to group three or four are gathered into cluster III, and the remaining 9-*cis*-RA low-energy conformations are found in clusters IV and V.

Estimating RA Conformational Energies with M06 and M06-2x Functionals. Our earlier work showed that the M06 class of density functionals gave relative energies that reliably matched MP2/CBS calculations for Ala_3 done at HF/6-31G* optimized geometries.²¹ To further examine the performance of the M06 functional class, we performed a series of M06 and M06-2x single point energy calculations on the global and local minima of ATRA and 9-*cis*-RA. Since we are not particularly interested in highly distorted retinoic acid structures, only conformers placed into groups one, two, three, and four were selected as the starting geometries for both ATRA and 9-*cis*-RA, respectively (see Figures 3 and 5 for backbone conformations). Figure 6A reveals the relative energy values calculated for ATRA low-energy conformers at different levels of theory: HF/6-31+G(d), M06/6-311+G(d,p), M06/aug-cc-pVDZ, MP2/6-311+G(d,p), MP2/aug-cc-pVDZ, MP2/aug-cc-pVTZ, and MP2/CBS. HF/6-31+G(d) underestimated the relative energies of the conformers in ATRA groups two, three, and four (see REA05 to REA16 in Figure 6A) by about 2 kcal/mol, notwithstanding its performance around the global minimum was acceptable (REA01 to REA04, Figure 6A). For single point calculations with the M06 functional, the relative

energy provided by the 6-311+G(d,p) basis set overestimated the MP2/CBS results by about 0.6 kcal/mol. M06/aug-cc-pVDZ performs better than M06/6-311+G(d,p) yet its relative energies are at least 0.3 kcal/mol higher than the MP2/CBS values (Figure 6A). Figure 6C shows a similar outcome for the M06 functional for the 9-*cis*-RA minimum energy conformers. Note that both Figure 6A and C use REA01 to REA16 along the x-axis labels, yet they represent entirely different sets of ATRA and 9-*cis*-RA conformations. Interestingly, a comparison of the M06-2x and MP2 methods combined with the 6-311+G(d,p), aug-cc-pVDZ, and aug-cc-pVTZ basis sets (see Figure 6B) shows excellent agreement. All three M06-2x relative energy curves essentially coincide with the MP2/CBS results (Figure 6B). The same trend is repeated for the 16 9-*cis*-RA low-energy conformations (see Figure 6D). The M06/aug-cc-pVTZ calculation data was not shown due to SCF convergence issues (even if we set the SCF convergence criterion to 10^{-4} au, the energies still cannot converge).

On the basis of the excellent match between the M06-2x and MP2/CBS curves, we were also interested in quantitatively determining whether any of the DFT methods employed (M06-2x/6-311+G(d,p), M06-2x/aug-cc-pVDZ, and M06-2x/aug-cc-pVTZ) differed from MP2/CBS in terms of the calculated RA conformational energies. If we regarded the RA low-energy conformers as 16 blocks and these quantum chemistry methods as four different treatments, this question can be answered by using distribution-free two-side all-treatment multiple comparisons based on Friedman rank sums (Wilcoxon–Nemenyi–McDonald–Thompson test), since the distribution of RA conformational energies was unknown and might not be a normal distribution (theoretically the conformational energy should obey a Boltzmann distribution). From Appendix I in the Supporting Information, we can conclude that M06-2x/aug-cc-pVDZ, M06-2x/aug-cc-pVTZ, and MP2/CBS are equivalent with respect to the calculated conformational energies for both ATRA and 9-*cis*-RA (failure to reject the null hypothesis at a *p* value > 0.05 for each pair comparison). We cannot find strong statistical evidence to claim that M06-2x/6-311+G(d,p) is not

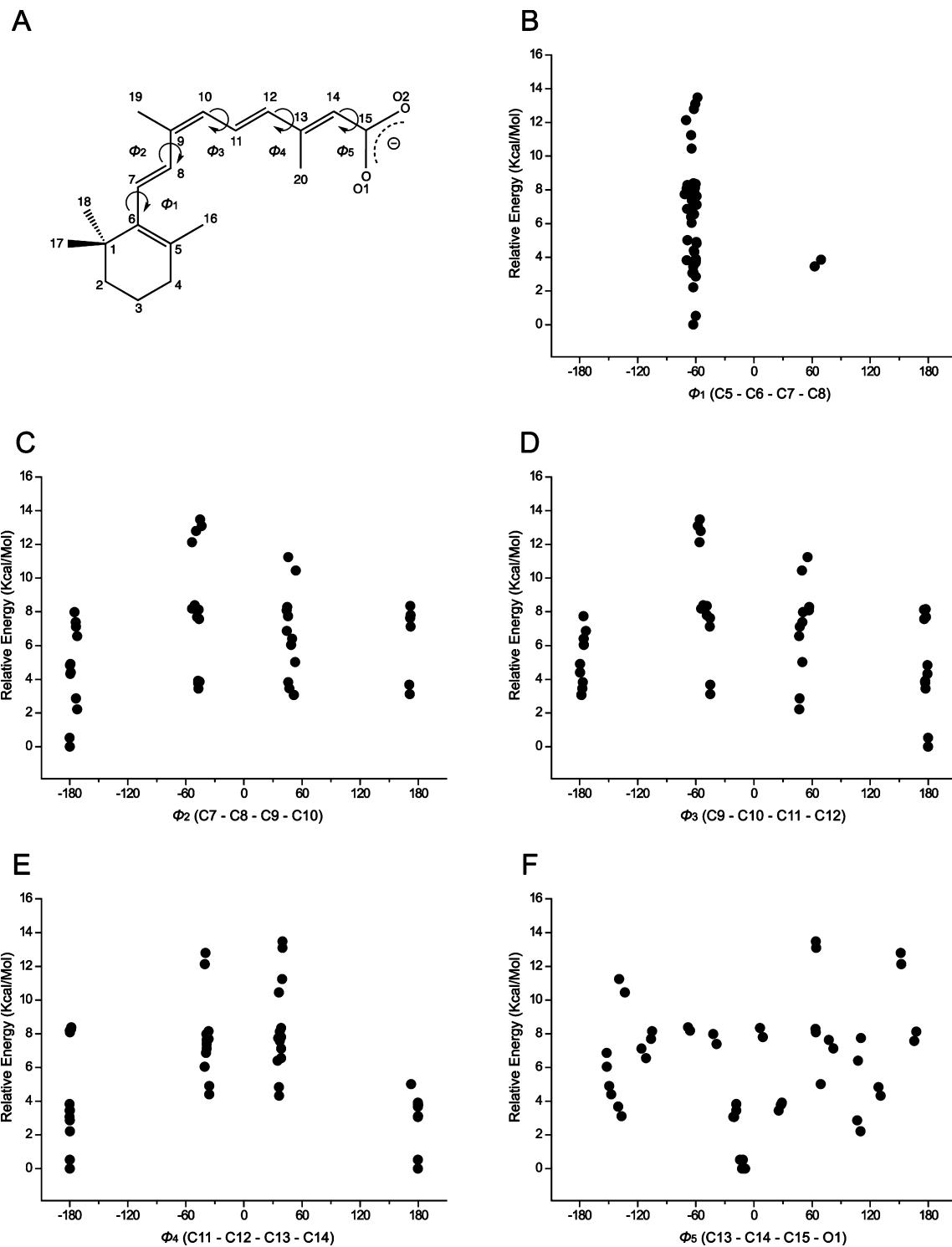


Figure 4. (A) Structure of 9-cis retinoic acid (9-cis-RA). The backbone torsion angles are indicated for the 9-cis-RA. (B–F) Relationship between relative conformational energies of the 48 9-cis-RA minimum energy conformations and five selected dihedral angles (in degrees) along the polyene chain. Relative conformational energies were calculated at the MP2/aug-cc-pVDZ level in the gas phase.

different from MP2/CBS for either ATRA or 9-cis-RA (reject the null hypothesis in favor of alternative hypotheses at a p value < 0.05). Overall, the single-point energy results indicate that the M06-2x functional is suitable for predicting the relative conformational energies of retinoic acid in the gas phase when larger basis sets are utilized.

Conformational Preferences of Protein Bound All-trans and 9-cis-RAs. Having probed the conformational

preferences of ATRA and 9-cis-RA in the free state, it is illustrative to compare and contrast with the conformations of RA in target protein binding sites. Previous studies in our lab and others have already demonstrated that a QM/MM X-ray refinement approach is capable of enhancing the quality of X-ray structures of metalloproteins and receptors with bound drug-like molecules.^{18,22–31} Therefore, we selected four ATRA and one 9-cis-RA–protein complex crystal structure from the

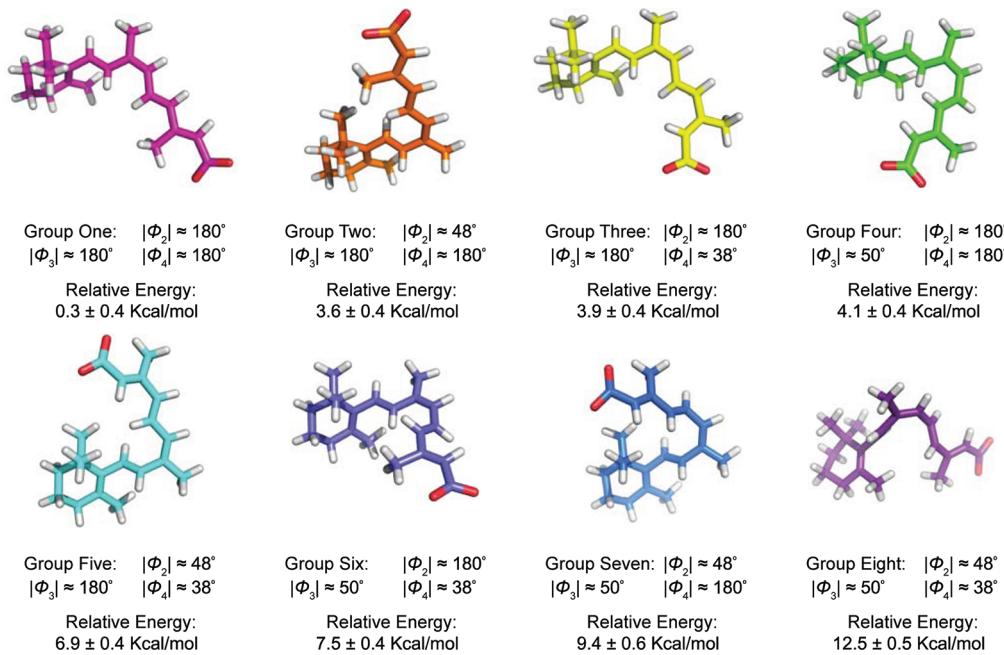


Figure 5. Eight groups of energy preferred conformers characterized for 9-cis-RA arranged in order of increasing average relative conformational energy.

PDB database and carried out hybrid QM/MM X-ray refinement on each of them (see Materials and Methods section and Li et al. for detail¹⁸). This study is an extension of our recent work on ATRA QM/MM refinement, but herein we largely focus on the comparison between the refined bioactive conformers and the global minimum of free retinoic acid. Figure 7 shows the overlays of the initial PDB structures of ATRA and 9-cis-RA with the corresponding refined structures. It is clear that after QM/MM X-ray refinement every bound retinoic acid conformer experienced changes in their spatial orientation and position. Table 1 lists the backbone Φ_1 to Φ_5 torsion angles (absolute values) of the six selected bound retinoic acid conformations as well as their QM/MM refined structures. In addition to this analysis, we also measured the dihedral angles between two planes whose line of intersection contains a methyl-substituted double bond (ψ_1 (C16–C5–C6–C7), ψ_2 (C19–C9–C10–C11), and ψ_3 (C20–C13–C14–C15), see Table 1). From Figure 7A and Table 1, we can see that C19 moves in the refined 3CWK ATRA conformer, accompanied by a torsion angle ψ_2 (C19–C9–C10–C11) change from 62° to 2° (the free value found above should be either 0 or 180°). When it comes to 1CBS, no significant changes took place in the polyene backbone, and the most remarkable displacement caused by the QM/MM refinement was involved in the rotation of the C6–C7 single bond, Φ_1 (C5–C6–C7–C8), which adjusts from 33° to 42° and is much closer to the favorable value for this torsion angle ($\approx 63^\circ$, see Figure 7B and Table 1). Similarly, the 2LBD and its refined structure are quite similar with the Φ_1 (C5–C6–C7–C8) angle experiencing the largest change from 45° to 48° (Figure 7C and Table 1). Furthermore, we compared the refined structures with the low energy free RA conformations using Tanimoto scores as implemented in ROCS. Indeed, the QM/MM refined RA conformations are more like the local and global minima than their corresponding original PDB geometries, which is evident from the heatmap shown as Supporting Information Figure 2.

Former studies in our lab have already demonstrated that a QM/MM X-ray refinement approach was a potent tool in fixing the internal coordinate errors of drug-like molecules bound to their receptors, which leads to an immense reduction of ligand conformational strain energy.³² In other words, strain energy itself could be employed as a criterion to quantitatively evaluate the effect of our QM/MM X-ray refinement method. Therefore, we extracted the coordinates of the retinoic acid conformers from each selected PDB RA–protein complex, as well as the refined structure, and computed the global and local strain energy in both the gas phase and PCM solvent model. Here, we define the global strain energy as the energy difference between the crystallographic conformation and the lowest energy conformation of the free ligand, whereas the local strain energy of a bound ligand represents the energy difference between the experimentally bound conformation and the nearest local minimum on conformational energy surface obtained by unconstrained geometry optimization. A similar strategy was used to look for the ATAR and 9-cis-RA global minimum in an aqueous environment, which means we carried out the HF/6-31G(d) geometry optimization on all Omega identified low energy conformers with the PCM solvent model, and the results are shown in Supporting Information Figure 3.

Table 2 gives the collected data for the strain energy calculations in both the gas phase and PCM solvent model using the MP2/CBS method. In the case of 3CWK, the optimization of the entire geometry substantially shifts the C19 atom, causing a 97.1% reduction in the local strain energy ($1 - 1.9/65.1 = 0.971$, see Table 2). For 1CBS and 2LBD, even though only a few torsion angles were moderately optimized via QM/MM refinement, the local strain dropped from 22.4 to 3.0 and 30.4 to 7.0, respectively (Table 2). Both the 2VE3 A1445 and B1445 ATRA PDB deposited conformations had distorted polyene chains, and after QM/MM X-ray refinement, an apparent change of the backbone was seen in each structure (Figure 7D and E). The Φ_2 and Φ_3 values of the 2VE3 experimental ATRA conformations are 163° and 155° for A1445 and

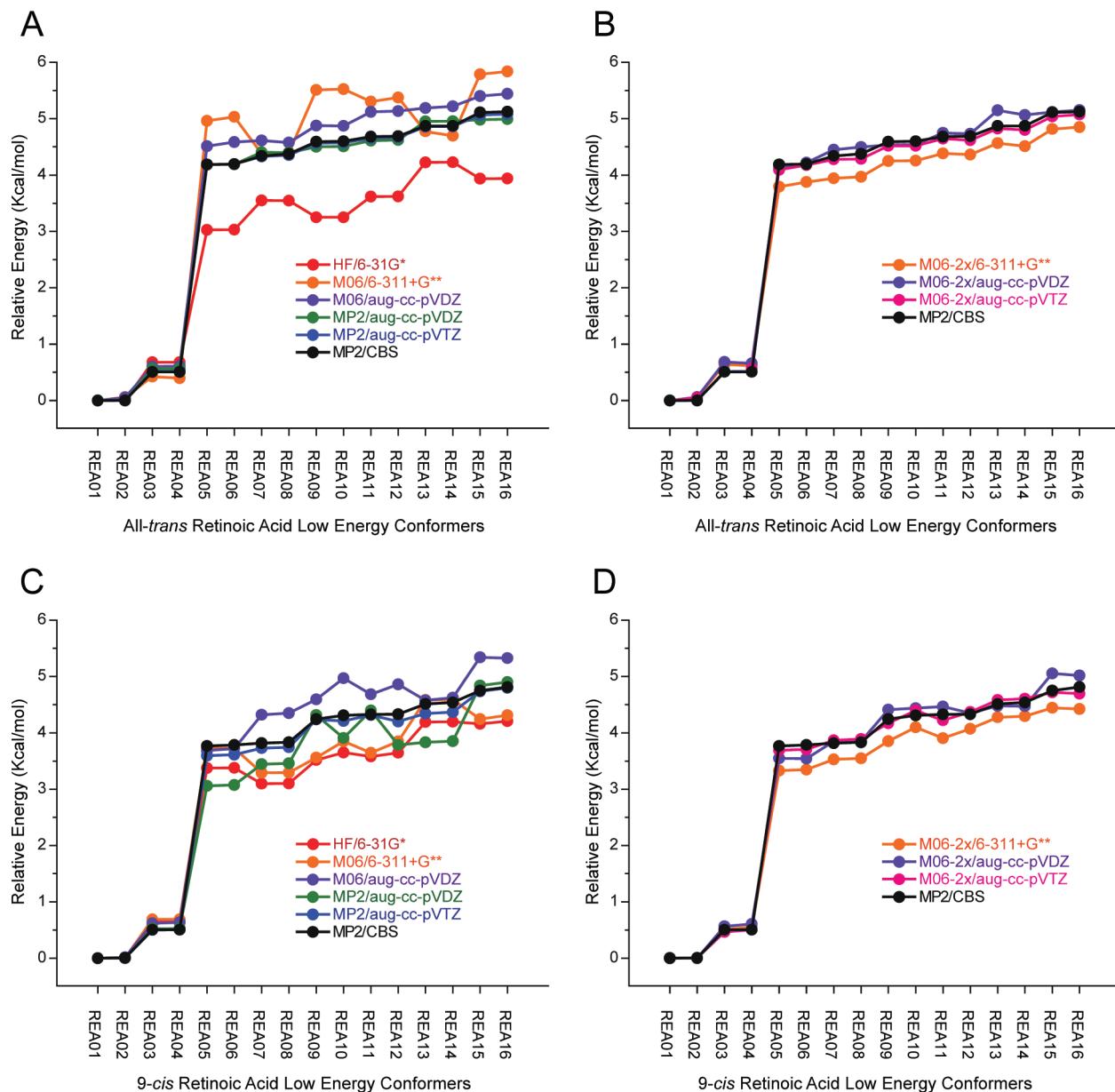


Figure 6. The computed relative energies of retinoic acid conformers obtained using (A) M06 and MP2 for ATRA, (B) M06-2x and MP2 for ATRA, (C) M06 and MP2 for 9-cis-RA, and (D) M06-2x and MP2 for 9-cis-RA with different basis sets in the gas phase.

159° and 161° for B1445, while after QM/MM refinement, they are 178° and 172° for A1445 and 178° and 179° for B1445 (see Table 1), which are in better accord with the low energy conformers observed in the gas and PCM solution phase ($\approx 180^\circ$). Moreover, the positions of the C19 and C20 atoms were poorly resolved, which distorted these two methyl groups and twisted the ψ_2 (C19–C9–C10–C11) = 20° and ψ_3 (C20–C13–C14–C15) = 28° for A1445 and ψ_2 (C19–C9–C10–C11) = 13° and ψ_3 (C20–C13–C14–C15) = 28° for B1445, respectively. QM/MM refinement relaxed the ψ_2 torsion to 2° and ψ_3 to 5° for A1445 as well as adjusting the ψ_2 torsion to 6° and ψ_3 to 11° for B1445. The Φ_4 (C11–C12–C13–C14) torsion angles for both the A1445 and B1445 PDB deposited conformers remained distorted even after QM/MM refinement, and this is the main reason why the rerefined A1445 and B1445 ATRA structures still have relatively high local strain energies (6.7 kcal/mol for A1445 and 7.0 kcal/mol for B1445). However, in comparison with the

115.0 and 101.3 kcal/mol conformational strain energies found in the initial PDB structures, our QM/MM refinement results for 2VE3 are acceptable. Ultimately, although the 1XDK crystallographic and QM/MM refined conformations are nearly identical (Table 1 and Figure 7F), the local strain trends downward from 40.0 to 7.1 kcal/mol (Table 2). Strain energies calculated by other levels of theory other than MP2/CBS are presented in Table 1 in the Supporting Information.

Table 2 also reveals the role that solvent plays in the calculation of strain energy. In the gas phase, the local strain energy of the PDB deposited coordinates ranged from 22.4 to 115.0 kcal/mol with a mean of 62.4 kcal/mol, while the PCM solvent model reduced the range of local strain energies from 19.5 to 112.2 kcal/mol with an average value of 60.1 kcal/mol. The strain energies calculated for the crystallographically determined ligands relative to the global minimum ranged from 22.9 to 115.0 kcal/mol *in vacuo* and 19.8 to 112.3 kcal/mol in aqueous solution. For the QM/MM refined structures,

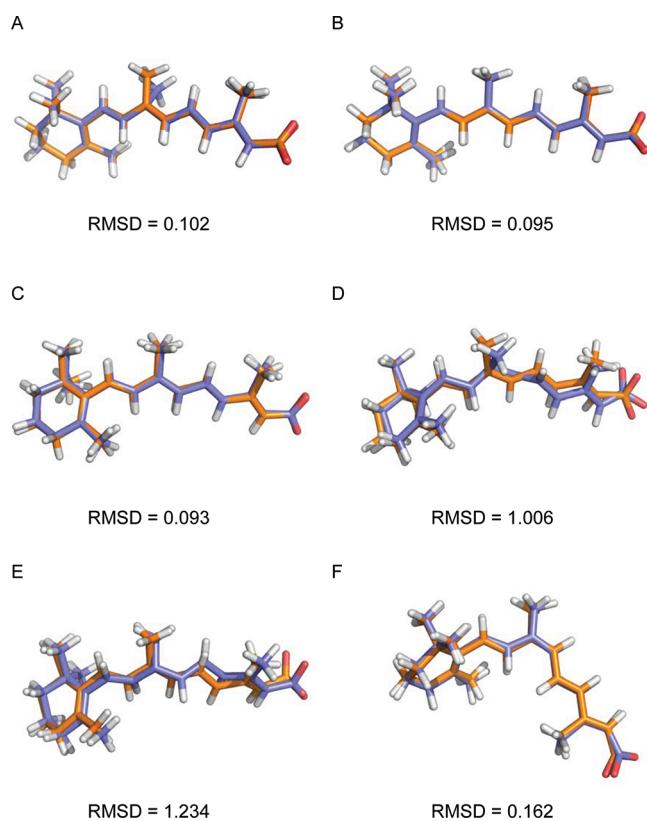


Figure 7. Overlay of the PDB deposited retinoic acid conformers (purple) and their corresponding QM/MM rerefined structures (orange). (A) 3CWK A300, (B) 1CBS A200, (C) 2LBD A500, (D) 2VE3 A1445, (E) 2VE3 B1445, (F) 1XDK A500.

the local strain energy ranged from 1.9 to 7.1 kcal/mol with an average value of 5.4 kcal/mol in the gas phase. Local strain calculations using a continuum solvent model ranged from 2.1 to 7.3 for refined structures with a mean value of 4.7 kcal/mol. Similarly, the PCM solvent model gave a decreased global strain energy ranging from 2.3 to 7.4 kcal/mol versus 2.4 to 7.6 kcal/mol for the refined retinoic acid conformations in the gas phase. Note the variability of solvation effects on the different refined structure. The 1XDK 9-cis-RA refined structure has the most significant reduction of local strain energy in aqueous solution ($7.1 - 5.1 = 2.0$ kcal/mol) followed by 2LBD ($7.0 - 5.6 = 1.4$ kcal/mol) and then the 1CBS and 2VE3 A1445 refined conformers ($3.0 - 2.1 = 0.9$ kcal/mol and $6.7 - 6.2 = 0.5$ kcal/mol, respectively). However compared to the results in the gas phase, the 3CWK and 2VE3 B1445 refined conformations

underwent a slight increase in local strain energy with the PCM solvent model, mainly due to the difference between the local minimum identified in the gas phase and in the PCM solvent model. In general, solvent had a stabilizing effect for both original PDB and refined RA structures, which is reflected by the overall decline of local and global strain energies. Note that Table 3 reports the real-space *R* factor at a given weighting factor, from which we deduce that the reduction in strain energies was not caused by overfitting to quantum mechanics (the weighting factors are small (1 or less) and the real-space *R* factors are ~ 0.1 or less). Table 3 also shows that our QM/MM refinement protocol yields to lower real-space *R* factors than the conventional CNS method (force field only), which demonstrates that the QM/MM rerefined RA conformers better fit the experimental electron density data. The RMSD values in Figure 7 also reveal that strain relief arises due to small changes in the geometries between the original PDB and the rerefined RA conformers.

CONCLUSIONS

The present study describes a thorough conformational analysis of retinoic acid using Omega and quantum mechanical calculations. A total of 66 and 48 energy minima have been identified and characterized by means of HF/6-31G(d) geometry optimization in the gas phase for ATRA and 9-cis-RA, respectively. It was found that the torsion angles associated with the ployene chain exert the most influence on the conformational energy, which in turn affects the retinoic acid backbone orientation. On the basis of the absolute values of these torsion angles, we divided the 66 ATRA low energy conformers into eight groups. The low-energy ATRA conformers, not unexpectedly, prefer the double bonds to be coplanar, which is reflected by the reduction of the relative potential energy. A similar structure–energy relationship was also observed in the 48 9-cis-RA low energy conformations. Additionally, the single-point energies using the M06 and M06-2x functionals with different basis sets were calculated. The result of these calculations indicates that the M06-2x level of theory with the aug-cc-pVDZ or aug-cc-pVTZ basis set is satisfactory in estimating the conformational energy of both ATRA and 9-cis-RA at the HF/6-31G(d) optimized geometries in the gas phase.

Five selected ATRA conformers from four ATRA-protein crystal structures found in the PDB were previously investigated¹⁸ using the *ab initio* QM/MM X-ray refinement approach and are further analyzed herein with respect to the uncomplexed gas and solution phase conformers. In addition, we performed the same analysis on a freshly rerefined 9-cis-RA-RXR α complex (PDB ID: 1XDK) in this paper as well. Every

Table 1. Backbone Torsion Angles (in degrees) with Respect to the PDB Deposited Retinoic Acid Conformers and Their QM/MM Rerefined Geometries

	3CWK A300		1CBS A200		2LBD A500		2VE3 A1445		2VE3 B1445		1XDK A500	
	PDB	QM/MM	PDB	QM/MM	PDB	QM/MM	PDB	QM/MM	PDB	QM/MM	PDB	QM/MM
$ \Phi_1 (C5-C6-C7-C8)$	48.7°	49.9°	32.5°	41.7°	45.1°	48.1°	68.1°	68.5°	77.6°	64.3°	58.3°	45.4°
$ \Phi_2 (C7-C8-C9-C10)$	174.1°	172.0°	178.5°	169.8°	179.8°	179.7°	163.3°	178.7°	159.6°	178.3°	179.3°	160.2°
$ \Phi_3 (C9-C10-C11-C12)$	179.4°	178.6°	178.7°	176.1°	179.4°	178.8°	155.4°	172.4°	161.2°	179.3°	175.5°	174.2°
$ \Phi_4 (C11-C12-C13-C14)$	171.4°	173.7°	179.8°	178.9°	176.8°	177.2°	166.9°	160.8°	174.3°	150.8°	176.5°	169.2°
$ \Phi_5 (C13-C14-C15-O1)$	27.2°	24.8°	129.1°	49.0°	3.5°	8.7°	134.1°	151.5°	3.8°	126.1°	141.1°	142.4°
$ \psi_1 (C16-C5-C6-C7)$	1.8°	0.1°	6.2°	3.2°	2.2°	0.0°	1.4°	3.7°	0.3°	2.6°	0.2°	2.0°
$ \psi_2 (C19-C9-C10-C11)$	61.5°	2.4°	3.0°	3.3°	1.4°	7.0°	20.1°	2.1°	13.6°	6.1°	179.9°	173.2°
$ \psi_3 (C20-C13-C14-C15)$	4.9°	0.2°	0.2°	1.7°	1.3°	13.6°	27.7°	5.7°	28.3°	11.3°	0.5°	8.5°

Table 2. Calculated Strain Energies Using the MP2/CBS Method for Retinoic Acid Conformers Deposited in PDB and Their Corresponding QM/MM Rerefined Structures

		3CWK A300		1CBS A200		2LBD A500		2VE3 A1445		2VE3 B1445		1XDK A500	
		PDB	QM/MM	PDB	QM/MM	PDB	QM/MM	PDB	QM/MM	PDB	QM/MM	PDB	QM/MM
GAS	local strain	65.1	1.9	22.4	3.0	30.4	7.0	115.0	6.7	101.3	7.0	40.0	7.1
	global strain	65.6	2.4	22.9	3.5	30.4	7.0	115.0	6.7	101.3	7.0	40.5	7.6
PCM	local strain	64.2	2.1	19.5	2.1	27.7	5.6	112.2	6.2	99.6	7.3	37.1	5.1
	global strain	64.5	2.4	19.8	2.3	27.8	5.6	112.3	6.3	99.7	7.4	37.5	5.4

Table 3. The Weighting Factors, Real-Space R Factors (RSR), and Real-Space Correlation Coefficients (RSCC) for QM/MM Rerefined RA-Protein Complexes^a

	3CWK A300	1CBS A200	2LBD A500	2VE3 A1445	2VE3 B1445	1XDK A500
weighting factors (w_a)	0.3	0.4	1.0	1.0	1.0	1.0
RSR	0.049 (0.060)	0.084 (0.086)	0.102 (0.141)	0.072 (0.143)	0.076 (0.129)	0.030 (0.043)
RSCC	0.951 (0.940)	0.916 (0.914)	0.898 (0.859)	0.928 (0.857)	0.924 (0.871)	0.970 (0.957)

^aValues within parentheses are calculated RSR and RSCC using the CNS refinement at the same weighting factor.

rerefined retinoic acid conformation preferred to arrange all conjugated double bonds of the polyene tail in the same plane with Φ_1 (C5–C6–C7–C8) at the most favorable value ($\approx 63^\circ$), which is in good agreement with the conformational analysis results of free RA using Omega plus HF/6-31G(d) geometry optimization. As a consequence, the internal ligand strain energies of QM/MM refined RA structures are lower relative to those from the corresponding PDB deposited conformations. In the gas phase, the local strain energy computed at the MP2/CBS level of theory has an average value of 5.4 kcal/mol among refined RA conformers versus the 62.4 kcal/mol found for the original PDB structures. Local strain gauged relative to the PCM solvent model has an average value of 4.7 kcal/mol, while for the PDB conformers this mean value is 60.1 kcal/mol (relative to MP2/CBS results). The same decrease in ligand strain after QM/MM refinement was also detected in terms of the global strain energy. This again illustrates that QM/MM X-ray refinement is a powerful tool in improving the intrinsic coordinate errors of bound drug-like molecules, which are mainly induced by inadequate MM models used during standard X-ray refinement.

Native retinoic acid itself was not a viable oral drug molecule due to its unsatisfactory side effects, yet its derivatives, retinoids, and rexinoids such as TTAB (CD367),^{33,34} TTNPB (Ro 13–7419),^{33,35} TTNN,^{33,36} and LGD1069 (Bexarotene, Targretin; Ligand Pharmaceuticals)³⁷ play a critical role in anticancer treatment. The scheme for the rational design of retinoids/rexinoids focused on narrowing the conformational flexibility by locking one or more of the conjugated double bonds of the RA backbone, which in turn will produce an energetic benefit on binding to the target protein. In order to simplify the synthesis of RA derivatives, medicinal chemists have used benzene rings to lock Φ_1 (C5–C6–C7–C8), thereby forcing it to be 180° . However, according to our conformational analysis of both free and bound RA conformers, the low energy value of Φ_1 (C5–C6–C7–C8) is around 63° , implying that nonrigid rings might be better choices than the phenyl group. Hence, our present study is not merely some benchmark calculation for different computational chemistry methods, but it also gives insight into the discovery of more effective anticancer drugs.

MATERIALS AND METHODS

1. Generating Retinoic Acid Conformation Ensembles.

Two sets of retinoic acid geometries (one for all-*trans* and another for 9-*cis*) were generated with Omega version 2.4.3 from Openeye Scientific Software.³⁸ For the adjustable parameters “-buildff” and “-searchff”, we selected MMFF94s rather than the default MMFF94s_NoEstat (without Coulombic terms) force field since in our study the deprotonated form of retinoic acid was chosen for all calculations. The maximum allowed internal energy gap from the lowest energy conformer was set to 200 kcal/mol (“-ewindow 200”). Parameters “-maxconf” and “-rms” were both set to zero, and Omega does not take duplicate conformers into account and outputs all generated conformers. Consequently, 948 ATRA conformers and 752 9-*cis* conformers were generated by OMEGA. Further geometry optimization was done at the HF/6-31G* level using the Omega generated geometries as input. The HF/6-31G* level of theory was chosen for optimization to match the model chemistry and basis set used in our QM/MM X-ray refinement efforts. All *ab initio* geometry optimizations were done using the Gaussian 09 suite of programs³⁹ with the keyword “Opt = VeryTight”.

2. Single Point Energy Calculation. Single-point calculations were carried out at the HF/6-31G*, M06/6-311+G**, M06/aug-cc-pVDZ, M06-2x/6-311+G**, M06-2x/aug-cc-pVDZ, and M06-2x/aug-cc-pVTZ levels with Gaussian 09³⁹ in a vacuum and in the PCM solvent model. Appendix II in the Supporting Information gives the parameters used in the PCM calculations. The MP2/CBS method was applied to compare *ab initio* and DFT single-point energy calculations. The CBS limit was determined by extrapolating MP2/aug-cc-pVXZ for X = D and T, for each retinoic acid conformer using a two-parameter exponential function. For more details regarding the extrapolation scheme, please see our previous work on the conformational analysis of ibuprofen.³²

3. Alignment and Clustering of Retinoic Acid Conformers. The alignment of different retinoic acid conformers was accomplished using ROCS (Rapid Overlay of Chemical Structures) version 3.1.0.³⁸ All parameters were set to their default values. In this study, we used the ShapeTanimoto score to estimate the 3D structure similarity between two retinoic acid conformers. The value of the ShapeTanimoto score ranges from 0 to 1, where 1 indicates that the two molecules are

identical and 0 means their volumes (shapes) do not overlap. The ShapeTanimoto scores of the HF/6-31G* optimized Omega generated conformers were clustered with the centroid linkage method. Uncentered correlation was chosen as the similarity metric. Cluster 3.0 (using C clustering library version 1.5) was employed for the clustering calculation,⁴⁰ and the clustering results were plotted by using Java Treeview version 1.1.5r2.⁴¹

4. Quantum Refinement. The deposited PDB coordinates of 3CKW,⁴² 1CBS,⁴³ 2VE3,⁴⁴ 2LBD⁴⁵ (RAR γ), and 1XDK⁴⁶ (RXR α) served as the starting geometries for the QM/MM refinement procedure. In the QM/MM X-ray crystal structure refinement, our goal is to find out the minimum of the energy function:^{23–26,28–30,32,47–50}

$$E = E_{\text{QM/MM}} + w_a E_{\text{X-ray}}$$

where E is a pseudo-energy function where $E_{\text{QM/MM}}$ is the energy restraint obtained from the QM/MM calculation and $E_{\text{X-ray}}$ is a pseudo-energy penalty function that represents the difference between the observed and calculated X-ray structure factors. w_a is the weighting factor that balance the relative importance of $E_{\text{QM/MM}}$ and $E_{\text{X-ray}}$ on the total energy. In these calculations, the QM region covers the retinoic acid and residues and/or water molecules within 2 Å of the ligand itself. See Table 2 in the Supporting Information for the residues included in the QM regions for each RA–protein complex. For both chain A and chain B in cyanobacterial cytochrome P450 CYP120A1 (PDB entry 2VE3), the heme molecules were excluded from the QM region. The HF/6-31G* level of theory was selected as the QM Hamiltonian for the QM region. The SANDER program in Amber 10⁵¹ was the interface that merged the X-ray gradients with the gradients from the QM/MM calculation using Amber 10 plus our in-house *ab initio* code QUICK. The $E_{\text{X-ray}}$ value was obtained from the CNS (Crystallography and NMR System) package.⁵²

5. Nonparametric Test. All nonparametric statistical inferences were performed using R 2.14.1 with the “pgirmess” package.⁵³

■ ASSOCIATED CONTENT

Supporting Information

Clustering Tanimoto scores of ATRA and 9-*cis*-RA energy minimum conformations. 3D structure comparison between Omega/*ab initio* QM method identified retinoic acid local/global minimum and QM/MM refined X-ray structures. Conformational search of ATRA and 9-*cis*-RA low energy structures in PCM solvent model. Calculated RA strain energies via various quantum methods. Residues included in the QM region during refinement process. Details of nonparametric statistical inference. Parameters setting of PCM model. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 352-392-6973. Fax: 352-392-8722. E-mail: merz@qtp.ufl.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the NIH (SBIR GM079899 and RO1's GM044974 and GM066859) for financial support of this research. Computing support from the University of Florida High Performance Computing Center is gratefully acknowledged.

■ REFERENCES

- (1) de Lera, A. R.; Bourguet, W.; Altucci, L.; Gronemeyer, H. Design of selective nuclear receptor modulators: RAR and RXR as a case study. *Nat. Rev. Drug Discovery* **2007**, *6*, 811–820.
- (2) Tang, X. H.; Gudas, L. J. Retinoids, retinoic acid receptors, and cancer. *Annu. Rev. Pathol.: Mech. Dis.* **2011**, *6*, 345–364.
- (3) Tallman, M. S.; Andersen, J. W.; Schiffer, C. A.; Appelbaum, F. R.; Feusner, J. H.; Ogden, A.; Shepherd, L.; Willman, C.; Bloomfield, C. D.; Rowe, J. M.; Wiernik, P. H. All-trans-retinoic acid in acute promyelocytic leukemia. *N. Engl. J. Med.* **1997**, *337*, 1021–1028.
- (4) Warrell, R. P. Jr.; Frankel, S. R.; Miller, W. H. Jr.; Scheinberg, D. A.; Itri, L. M.; Hittelman, W. N.; Vyas, R.; Andreeff, M.; Tafuri, A.; Jakubowski, A.; et al. Differentiation therapy of acute promyelocytic leukemia with tretinoin (all-trans-retinoic acid). *N. Engl. J. Med.* **1991**, *324*, 1385–1393.
- (5) Huang, M. E.; Ye, Y. C.; Chen, S. R.; Chai, J. R.; Lu, J. X.; Zhoa, L.; Gu, L. J.; Wang, Z. Y. Use of All-Trans Retinoic Acid in the Treatment of Acute Promyelocytic Leukemia. *Blood* **1988**, *72*, 567–572.
- (6) Weis, K.; Rambaud, S.; Lavau, C.; Jansen, J.; Carvalho, T.; Carmofonseca, M.; Lamond, A.; Dejean, A. Retinoic Acid Regulates Aberrant Nuclear-Localization of Pml-Rar-Alpha in Acute Promyelocytic Leukemia-Cells. *Cell* **1994**, *76*, 345–356.
- (7) Lai, A. Y.; Wade, P. A. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nat. Rev. Cancer* **2011**, *11*, 588–596.
- (8) Elstner, E.; Muller, C.; Koshizuka, K.; Williamson, E. A.; Park, D.; Asou, H.; Shintaku, P.; Said, J. W.; Heber, D.; Koeffler, H. P. Ligands for peroxisome proliferator-activated receptor-gamma and retinoic acid receptor inhibit growth and induce apoptosis of human breast cancer cells in vitro and in BNX mice. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 8806–8811.
- (9) Anzano, M. A.; Byers, S. W.; Smith, J. M.; Peer, C. W.; Mullen, L. T.; Brown, C. C.; Roberts, A. B.; Sporn, M. B. Prevention of Breast-Cancer in the Rat with 9-Cis-Retinoic Acid as a Single-Agent and in Combination with Tamoxifen. *Cancer Res.* **1994**, *54*, 4614–4617.
- (10) White, K. P.; Hua, S. J.; Kittler, R. Genomic Antagonism between Retinoic Acid and Estrogen Signaling in Breast Cancer. *Cell* **2009**, *137*, 1259–1271.
- (11) Tamura, K.; Nakae, D.; Horiguchi, K.; Akai, H.; Kobayashi, Y.; Andoh, N.; Satoh, H.; Denda, A.; Tsujuchi, T.; Yoshiji, H.; Konishi, Y. Inhibition by N-(4-hydroxyphenyl)retinamide and all-trans-retinoic acid of exogenous and endogenous development of putative preneoplastic, glutathione S-transferase placental form-positive lesions in the livers of rats. *Carcinogenesis* **1997**, *18*, 2133–2141.
- (12) Colston, K. W.; Pettersson, F.; Dalglish, A. G.; Bissonnette, R. P. Retinoids cause apoptosis in pancreatic cancer cells via activation of RAR-gamma and altered expression of Bcl-2/Bax. *Br. J. Cancer* **2002**, *87*, 555–561.
- (13) Lammer, E. J.; Chen, D. T.; Hoar, R. M.; Agnish, N. D.; Benke, P. J.; Braun, J. T.; Curry, C. J.; Fernhoff, P. M.; Gris, A. W.; Lott, I. T.; Richard, J. M.; Sun, S. C. Retinoic Acid Embryopathy. *N. Engl. J. Med.* **1985**, *313*, 837–841.
- (14) Kessel, M.; Gruss, P. Homeotic Transformations of Murine Vertebrae and Concomitant Alteration of Hox Codes Induced by Retinoic Acid. *Cell* **1991**, *67*, 89–104.
- (15) Miller, W. H. The emerging role of retinoids and retinoic acid metabolism blocking agents in the treatment of cancer. *Cancer* **1998**, *83*, 1471–1482.
- (16) Beppu, Y.; Kakitani, T. Comprehensive Study of Molecular-Conformation and Adiabatic Potential Surface of All-Trans Retinal by the Cndo/2 Method. *Chem. Phys.* **1990**, *148*, 333–346.

- (17) vanAalten, D. M. F.; deGroot, B. L.; Berendsen, H. J. C.; Findlay, J. B. C. Conformational analysis of retinoids and restriction of their dynamics by retinoid-binding proteins. *Biochem. J.* **1996**, *319*, 543–550.
- (18) Li, X.; Fu, Z.; Merz, K. M. Jr. QM/MM refinement and analysis of protein bound retinoic acid. *J. Comput. Chem.* **2012**, *33*, 301–310.
- (19) Gilardi, R.; Karle, I. L.; Karle, J.; Sperling, W. Crystal structure of the visual chromophores, 11-cis and all-trans retinal. *Nature* **1971**, *232*, 187–189.
- (20) Klucik, J.; Xiao, Y. D.; Hammond, P. S.; Harris, R.; Schmitt, J. D. Targacept active conformation search: a new method for predicting the conformation of a ligand bound to its protein target. *J. Med. Chem.* **2004**, *47*, 6831–6839.
- (21) Merz, K. M.; Molnar, L. F.; He, X.; Wang, B. Further analysis and comparative study of intermolecular interactions using dimers from the S22 database. *J. Chem. Phys.* **2009**, *131*.
- (22) Ryde, U.; Nilsson, K. Quantum chemistry can locally improve protein crystal structures. *J. Am. Chem. Soc.* **2003**, *125*, 14232–14233.
- (23) Nilsson, K.; Ryde, U. Protonation status of metal-bound ligands can be determined by quantum refinement. *J. Inorg. Biochem.* **2004**, *98*, 1539–1546.
- (24) Yu, N.; Yennawar, H. P.; Merz, K. M. Jr. Refinement of protein crystal structures using energy restraints derived from linear-scaling quantum mechanics. *Acta Crystallogr., Sect. D* **2005**, *61*, 322–332.
- (25) Yu, N.; Hayik, S. A.; Wang, B.; Liao, N.; Reynolds, C. H.; Merz, K. M. Assigning the protonation states of the key aspartates in beta-Secretase using QM/MM X-ray structure refinement. *J. Chem. Theory Comput.* **2006**, *2*, 1057–1069.
- (26) Yu, N.; Li, X.; Cui, G.; Hayik, S. A.; Merz, K. M. II. Critical assessment of quantum mechanics based energy restraints in protein crystal structure refinement. *Protein Sci.* **2006**, *15*, 2773–2784.
- (27) Ryde, U.; Hsiao, Y. W.; Rulisek, L.; Solomon, E. I. Identification of the peroxy adduct in multicopper oxidases by a combination of computational chemistry and extended X-ray absorption fine-structure measurements. *J. Am. Chem. Soc.* **2007**, *129*, 726–727.
- (28) Li, X.; He, X.; Wang, B.; Merz, K. Jr. Conformational variability of benzamidinium-based inhibitors. *J. Am. Chem. Soc.* **2009**, *131*, 7742–7754.
- (29) Hsiao, Y. W.; Sanchez-Garcia, E.; Doerr, M.; Thiel, W. Quantum refinement of protein structures: implementation and application to the red fluorescent protein DsRed.M1. *J. Phys. Chem. B* **2010**, *114*, 15413–15423.
- (30) Li, X.; Hayik, S. A.; Merz, K. M. Jr. QM/MM X-ray refinement of zinc metalloenzymes. *J. Inorg. Biochem.* **2010**, *104*, 512–522.
- (31) Ryde, U.; Greco, C.; De Gioia, L. Quantum refinement of [FeFe] hydrogenase indicates a dithiomethylamine ligand. *J. Am. Chem. Soc.* **2010**, *132*, 4512–4513.
- (32) Fu, Z.; Li, X.; Merz, K. M. Jr. Accurate assessment of the strain energy in a protein-bound drug using QM/MM X-ray refinement and converged quantum chemistry. *J. Comput. Chem.* **2011**, *32*, 2587–2597.
- (33) Bollag, W. Arotinoids. A new class of retinoids with activities in oncology and dermatology. *Cancer Chemother. Pharmacol.* **1981**, *7*, 27–29.
- (34) Charpentier, B.; Bernardon, J. M.; Eustache, J.; Millois, C.; Martin, B.; Michel, S.; Shroot, B. Synthesis, structure-affinity relationships, and biological activities of ligands binding to retinoic acid receptor subtypes. *J. Med. Chem.* **1995**, *38*, 4993–5006.
- (35) Westin, S.; Kurokawa, R.; Nolte, R. T.; Wisely, G. B.; McInerney, E. M.; Rose, D. W.; Milburn, M. V.; Rosenfeld, M. G.; Glass, C. K. Interactions controlling the assembly of nuclear-receptor heterodimers and co-activators. *Nature* **1998**, *395*, 199–202.
- (36) Harant, H.; Korschineck, I.; Krupitza, G.; Fazeny, B.; Dittrich, C.; Grunt, T. W. Retinoic acid receptors in retinoid responsive ovarian cancer cell lines detected by polymerase chain reaction following reverse transcription. *Br. J. Cancer.* **1993**, *68*, 530–536.
- (37) Gottardis, M. M.; Bischoff, E. D.; Shirley, M. A.; Wagoner, M. A.; Lamph, W. W.; Heyman, R. A. Chemoprevention of mammary carcinoma by LGD1069 (Targretin): an RXR-selective ligand. *Cancer Res.* **1996**, *56*, 5566–5570.
- (38) OEChem, version 1.7.4; OpenEye Scientific Software, Inc.: Santa Fe, NM, 2010.
- (39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazayev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
- (40) de Hoon, M. J.; Imoto, S.; Nolan, J.; Miyano, S. Open source clustering software. *Bioinformatics* **2004**, *20*, 1453–1454.
- (41) Saldanha, A. J. Java Treeview--extensible visualization of microarray data. *Bioinformatics* **2004**, *20*, 3246–3248.
- (42) Geiger, J. H.; Vaezeslami, S.; Jia, X. F.; Vasileiou, C.; Borhan, B. Structural analysis of site-directed mutants of cellular retinoic acid-binding protein II addresses the relationship between structural integrity and ligand binding. *Acta Crystallogr., Sect. D* **2008**, *64*, 1228–1239.
- (43) Kleywegt, G. J.; Bergfors, T.; Senn, H.; Le Motte, P.; Gsell, B.; Shudo, K.; Jones, T. A. Crystal structures of cellular retinoic acid binding proteins I and II in complex with all-trans-retinoic acid and a synthetic retinoid. *Structure* **1994**, *2*, 1241–1258.
- (44) Kuhnel, K.; Ke, N.; Cryle, M. J.; Sligar, S. G.; Schuler, M. A.; Schlichting, I. Crystal structures of substrate-free and retinoic acid-bound cyanobacterial cytochrome P450 CYP120A1. *Biochemistry* **2008**, *47*, 6552–6559.
- (45) Renaud, J. P.; Rochel, N.; Ruff, M.; Vivat, V.; Chambon, P.; Gronemeyer, H.; Moras, D. Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature* **1995**, *378*, 681–689.
- (46) Bourguet, W.; Pogenberg, V.; Guichou, J. F.; Vivat-Hannah, V.; Kammerer, S.; Perez, E.; Germain, P.; de Lera, A. R.; Gronemeyer, H.; Royer, C. A. Characterization of the interaction between retinoic acid receptor/retinoid X receptor (RAR/RXR) heterodimers and transcriptional coactivators through structural and fluorescence anisotropy studies. *J. Biol. Chem.* **2005**, *280*, 1625–1633.
- (47) Jack, A. a. L. M. Refinement of large structures by simultaneous minimization of energy and R factor. *Acta Crystallogr., Sect. A* **1978**, *34*, 931–935.
- (48) Hendrickson, W. A. Stereochemically restrained refinement of macromolecular structures. *Methods Enzymol.* **1985**, *115*, 252–270.
- (49) Ryde, U.; Olsen, L.; Nilsson, K. Quantum chemical geometry optimizations in proteins using crystallographic raw data. *J. Comput. Chem.* **2002**, *23*, 1058–1070.
- (50) Nilsson, K.; Hersleth, H. P.; Rod, T. H.; Andersson, K. K.; Ryde, U. The protonation status of compound II in myoglobin, studied by a combination of experimental data and quantum chemical calculations: quantum refinement. *Biophys. J.* **2004**, *87*, 3437–3447.
- (51) Case, D. A.; Cheatham, C. L., III; Simmerling, J.; Wang, R. E.; Duke, R.; Luo, M.; Crowley, R. C.; Walker, W.; Zhang, K. M.; Merz, B.; Wang, S.; Hayik, A.; Roitberg, G.; Seabra, I.; Kolossaváry, K. F.; Wong, F.; Paesani, J.; Vanicek, X.; Wu, S. R.; Brozell, T.; Steinbrecher, H.; Gohlke, L.; Yang, C.; Tan, J.; Mongan, V.; Hornak, G.; Cui, D. H.; Mathews, M. G.; Seetin, C.; Sagui, V.; Kollman, P. A. AMBER 10; University of California: San Francisco, CA, 2008.
- (52) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grossé-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L.

Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr, Sect. D* **1998**, *54*, 905–921.

(S3) *R: A language and environment for statistical computing*; R Foundation for Statistical Computing: Vienna, Austria, 2011. ISBN 3-900051-07-0. <http://www.R-project.org/> (accessed March 2012).