# Predicting the Cell Differentiation Activity of 1α,25-Dihydroxyvitamin D₃ Side Chain Analogues from Docking Simulations

Fredy Sussman,* Antonio Rumbo,
M. Carmen Villaverde, and Antonio Mouriño

*Departamento de Química Orgánica y Unidad Asociada al CSIC, Facultad de Química, Universidad de Santiago de Compostela, Santiago de Compostela 15782, Spain*

*Received July 23, 2003*

**Abstract:** We present a receptor-based protocol for the prediction of the cell differentiation activities of a series of side chain analogues of 1α,25-dihydroxyvitamin D₃, a compound that exhibits a very large variety of biological functions. Our protocol is able to reproduce the activity of the compounds studied here. It also sheds light on the relative importance of binding site residues in the biological activity and on the mechanism behind it.

**Introduction.** The vitamin D receptor (VDR) is a ligand-dependent transcriptional regulator that belongs to the nuclear receptor (NR) transcription factor superfamily, a group of proteins that includes steroid, retinoid, and thyroid receptors as well as other receptors of unknown function.[1] The cognate ligands of these receptors act as molecular switches in gene transactivation by changing, upon binding, the conformation of the proteins to an active form that heterodimerizes with retinoic X receptor (RXR) (another member of the nuclear receptor family). The complex formed by these two proteins (VDR and RXR) binds to the direct repeat of the target gene promoter and then recruits coactivators.[2−4] The cognate ligand of VDR is the active form of vitamin D, 1α,25-dihydroxyvitamin D₃ [1α,25-(OH)₂D₃, **1**], a compound that has potent cell differentiation activity and shows immunosuppressive effects.[5] Most of the 1α,25-(OH)₂D₃ derivatives reported to date involve side chain modifications of the parent compound.[6] These derivatives were synthesized in order to take advantage of the highly diverse biological activities of **1** while trying to improve on its biological profile, which has been limited by hypercalcemic effects. The actual and potential therapeutic applications of these compounds range from antiproliferative diseases such as psoriasis, cancer, as well as bone disorders such as osteoporosis, and a variety of immune disorders.[7]

The search for drug leads could be empowered by computationally based methods that would predict the biological activity of compounds prior to synthesis. The earliest attempts to predict the activity of a wide range of vitamin D analogues involved the use of ligand-based protocols that relied on the conformational analysis of the analogue's side chain.[8,9] These studies led to the development of the active space group concept, which predicts that the most active ligands will be those whose side chains are l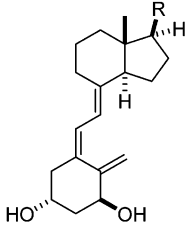ocalized in a specific conformational region, identified as EA.[9a] The structure determination of **1** and two of its 20-epi analogues (**4** and **7**), when bound to the ligand binding domain (LBD) of the VDR,[10,11] has paved the way for quantitative structure–function relationships.[12] In this paper we present a receptor-based protocol that relies on the analysis of models of the docked ligand structures with VDR. Starting from the known crystallographic structures, we have modeled the complexes between the LBD and a set of analogues with a variety of side chain motifs. The search for a predictive protocol for the activity induced by a ligand led us to evaluate and analyze the interaction energy of the ligands with various sets of residues. The residues in question range from those that form the LBD of the VDR to certain amino acid clusters that line up the VDR binding site. We chose this latter group of residues from those whose Ala scanning produced inactive VDR strains.[12] The work described here demonstrates that some estimates of the contacts between the ligands and sets of active site residues, as provided by the van der Waals contacts in the modeled docked structures, agree with the cell differentiation activity ranking of the analogues with respect to the vitamin D hormone. The method presented here has the potential to predict the relative importance of the residues on the VDR conformational changes that lead to cell differentiation.

**Methods.** The complexes between the VDR and a set of 11 ligands were modeled. The ligands included **1** and 10 side chain derivatives of different lengths and different epimeric forms and several side chain analogues with oxygen replacement at various positions (Table 1). The starting point for our calculations were the known three-dimensional structures of the LBD region of the VDR receptor bound to **1** and two analogues (see **4** and **7** in Table 1).[10,11] Two of these complexes (PDB entries 1IE8 and 1DB1) have structural gaps from residues A375 to A377. This segment was reconstituted by using the coordinates of the structure in the complex with **4** (entry 1IE9), which does have the coordinates for this fragment. Both the missing segment and the hydrogen atoms of the receptor were added using the Biopolymer module in InsightII.[13] All of the His residues of the receptor were chosen to be uncharged. Two of the His residues (His 305 and His 397) serve as the side chain anchor for the ligand by providing a hydrogen bond (HB) donor or by acting as an acceptor for the hydroxyl group of the side chain through their imidazole groups.[10,11] His 305 was selected as the HB acceptor for the side chain hydroxyl group, since the imidazole group of this residue is already involved as an HB donor for Gln 400;[10,11] His 397 was chosen as the HB donor (see Figure 1).

The docking process was performed in three stages:

**a. Conformation Generation.** The side chain variants of **1** were constructed with the help of the Builder module available in the InsightII suite of programs.[13] The initial conformations of the ligands were then energy-minimized, and their structure was superimposed onto **1**, ensuring that the hydroxyl group at the side chain of the analogue was placed as close as

* To whom correspondence should be addressed. Phone: 34 981 563 100. Fax: 34 981 591 014 or 34 981 595 012 (department). E-mail: fsussman@usc.es.

**Table 1.** Chemical Structures and Biological Activities of 1α,25-(OH)$_2$D$_3$ (**1**) and Its Side Chain Analogues

| No | R | A/A$_d$[a] |
|----|---|------|
| 1 | (structure, —OH) | |
| 2 | (structure, —OH) | 2-5[b] |
| 3 | (structure, O, —OH) | 15-30[c] |
| 4 | (structure, —OH) | 20-30[d] 100[e] |
| 5 | (structure, —OH) | 100[e] |
| 6 | (structure, —OH) | 200[d] |
| 7 | (structure, O, —OH) | 200000[d] |
| 8 | (structure, O, —OH) | 0.2[f] |
| 9 | (structure, O, —OH) | 20[d] |
| 10 | (structure, —OH) | 5[d] |
| 11 | (structure, O, —OH) | 5-10[g] |

[a] Relative activity referenced to **1**. [b] See refs 9a and 18. [c] See ref 19. [d] See ref 20. [e] See ref 21. [f] See ref 22. [g] See refs 9a and 23.



**Figure 1.** 1α,25-(OH)$_2$D$_3$ surrounded by some active site residues. These include the residues whose Ala mutants produce very low biological activity VDR strains and include His 397, a hydrogen bond donor to the side chain hydroxyl group (see Methods for details). For the sake of clarity we have displayed only the side chains of the residues and omitted the nonpolar hydrogens.

possible to that in **1**. This latter step was achieved through the use of the Search & Compare module in InsightII. A set of 1000 conformations was then generated for every ligand. This conformational set was generated by running molecular dynamics (MD) simulations for the ligand in vacuum at 300 K. During the course of the MD run, the rings, the triene system, and the oxygen atom of the side chain were kept fixed for all ligands.

**b. Conformation Docking.** The 1000 conformers generated above were docked into the target protein in the same unique orientation observed in the crystallographic structures.[10,11] The docked structures were generated by overlapping the analogue conformers onto the reference ligand (**1**) in the template and then forming new complexes by exchanging them. Whenever available, we used the models based on the protein crystallographic structures. For instance, the model for the binding of **1** was based on the crystallographic structure of 1DB1, while the structure of the bound ligand **4** was based on the PDB entry 1IE9. Most other receptor−ligand complexes do not have a crystallo-
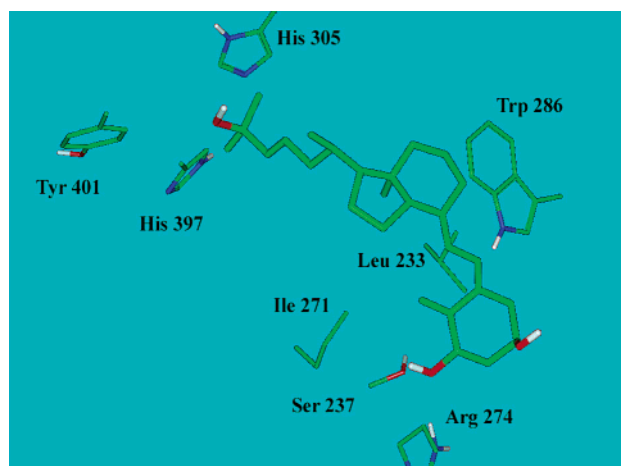
graphic structure; their structures were modeled on the basis of PDB entry 1IE9. The set of bound LBD structures were ranked in terms of the ligand−receptor interaction energy, after undergoing a very small energy minimization of 10 steps using the CHARMM force field.[14] The top 200 complexes were chosen from this ranking, and this subset was further optimized by a 1000-step steepest descent energy minimization for the whole system. This procedure affords some protein flexibility in our calculations.

**c. Activity Prediction.** We searched for a structure−activity relationship based on a possible correlation between the transcriptional activity and the interaction energy of the analogues with certain segments of the VDR. With this aim in mind, we analyzed the interaction energy of these derivatives with all the LBD residues, as well as with clusters of binding site residues. These amino acids were chosen from the Ala scanning work of Choi et al.[12] These authors performed Ala scanning of most of the VDR binding site residues and found that some of the Ala strains had an increased transcriptional activity while others retained or had markedly reduced biological potency when bound to some or all of the 1α,25-(OH)$_2$D$_3$ analogues studied in that work.[12] The strains chosen here have a significantly lower activity across the board irrespective of the ligand. The low-activity VDR strains include the L233A, S237A, I271A, R274A, W286A, H397A, and Y401A mutant strains. The interaction energies were evaluated using the CHARMM force field.[14] The placement of these residues around the ligand is represented in Figure 1.

**Results and Discussion.** The interaction energies of the 200 "best binding conformations" with the different sets of residues have been collected as energy distributions. The total interaction energy as well as its van der Waals component were evaluated for every set of receptor residues. The best correlation between the experimental activity and our calculations was obtained for the van der Waals component of the interaction energy with certain selected residues (see Figure 2). As one can see from Figure 2, the van der Waals contact energy represents a good index of activity. For instance,
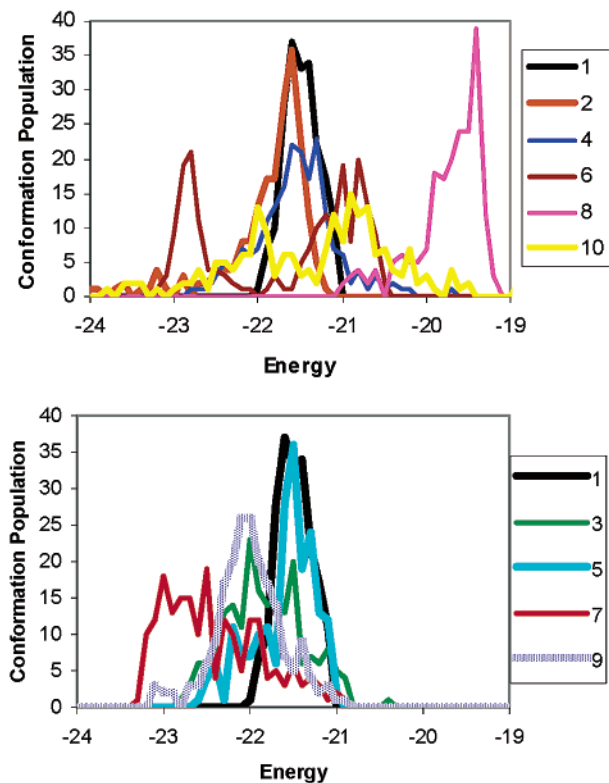
**Figure 2.** Energy distributions for the set of van der Waals interaction energies between the first 10 ligands, shown in Table 1, and some of the VDR binding site residues whose mutation by Ala abolishes transcriptional activity. The top panel displays the even numbered ligands, while the bottom panel displays the odd numbered ligands.

**Table 2.** Difference in the Interaction Energy Residue Contribution between Selected Analogues[a]

| residue | Δ(**1,7**) | Δ(**1,4**) | Δ(**11,8**) |
|---|---|---|---|
| all residues | −1.31 | −0.83 | 2.79 |
| Leu 233 | −0.28 | −0.37 | 0.05 |
| Ser 237 | −0.33 | −0.38 | 1.39 |
| Ile 271 | −0.78 | −0.72 | 0.34 |
| Arg 274 | 0.95 | 0.73 | 0.18 |
| Trp 286 | −0.80 | −0.63 | −0.34 |
| His 397 | 0.17 | 0.67 | 0.46 |
| Tyr 401 | −0.25 | −0.14 | 0.70 |

[a] The values listed are the difference in van der Waals contributions per residue for three pairs of ligands (**1** and **7**, **1** and **4**, and **8** and **11**), for the highest ranking conformer among the population of 200 conformers whose energy distribution is shown in Figure 2. The first entry row is the difference for all residues involved. Differences are taken between the last and the first entries in the argument of Δ. Energies are in kcal/mol.

our results indicate that compound **8** [23-oxa-1α,25-(OH)$_2$D$_3$] has a distribution located at the highest energy values of all the ligands studied, thus predicting the worst potency of all analogues shown in Table 1, an outcome in agreement with experimental results. On the other hand, the most potent analogues have the lowest energy distributions with their first peak located at lower values with respect to **1**, thus predicting a higher potency for these compounds, again in agreement with experimental results. The latter set of ligands include the superagonists, a set of compounds with cell differentiation activities that are at least 100−1000 times better than that of **1**. Establishing a correlation between the experimentally determined activity ratios and the ranking predicted by energy distribution for the higher potency ligands is hampered by the variability of the published experimental values for the activity ratio obtained for most ligands (see Table 1). The discrepancies arise from the diverse experimental protocols used, which often include different cell lines. Nevertheless, our protocol is capable of identifying the most potent ligands among the compounds studied here. The most active compounds are ligands **6** and **7** (known in the literature as ligands MC1301 and KH1060) with activities that are up to 5 orders of magnitude higher than that of **1** (see Table 1). It can be seen from Figure 2 that the first peak for the van der Waals interaction energy distribution for these two potent compounds lies at the lowest energy of all compounds studied. The only discrepancy between the observed activities and the contact energy index was found for ligand **11**. We have

found that this ligand's best binder conformers are localized at slightly higher energies than those of **1** (results not shown).

The present work differs from previous studies in that it is a receptor-based protocol that relies on the interaction energy of the ligands with a set of biologically important residues. One of the main features of this approach is the ability to predict the relative contribution of each residue to the biological activity induced by the ligands, a quantity that cannot be directly inferred from experimental work alone. This feature can be illustrated, for instance, by comparing the interaction energy profile of a single residue between any pair of ligands. To exemplify the single residue analysis capability of this protocol, we have chosen a more compact presentation. We have calculated the differences in single-residue contacts, between two analogues, for the highest ranking bound conformer among those shown in the conformation population displayed in Figure 2. The results are shown in Table 2. The analogues included in this table are **1**, **4**, **7**, **8**, and **11**. The difference between the residue contacts with ligands **4** and **1** was analyzed in order to determine the additional contacts, brought about by epimerization at C20, while the analysis of the difference between ligands **7** and **1** allows an insight into the effect of volume increase on activity. As seen from Table 1, ligands **4** and **7** differ in the size and chemical composition of the side chain. Nevertheless, these two ligands improve their predicted affinity by increasing their interaction energy with the same set of residues (see Table 2). Two of the residues that serve as anchoring points for the ligand (Arg 274 and His 397) have a decreased van der Waals interaction energy when the natural binder is replaced by the agonists **4** and **7**. Nevertheless, both epimer ligands reduce drastically their activity when bound to the Ala274 or Ala397 protein strains. This may indicate that the main role of these two residues is that of ligand structural anchors rather than molecular switches for activation. Finally, we have compared two analogues that differ only in having the oxygen displaced from O22 (analogue **11**) to O23 (analogue **8**). This simple change produces an analogue of volume similar to that of its parent compound but with the lowest activity of all compounds in Table 1. As seen from Table 2, analogue **8** owes its lower predicted activity (with respect to **11**) to a reduction of the interaction energy for all but one of the residues.

We have restricted our study to analogues that contain side chain modifications. Nevertheless, the changes on van der Waals interaction energy per residue are not restricted to the residues that contact the side chain but include the residues that contact the rings and the triene system (see Table 2). This may indicate that upon binding there is a degree of correlation between the displacements of the residues that contact directly with the side chain and those that interact with the rest of the ligand. The correlation shown here between the ratio of the cell differentiation activity and the location of the energy profile of the analogues contributed by some binding site residues may bear witness to the importance of the ligand binding step in the formation of a multicomponent functional complex of proteins and DNA. From our results it follows that the side chain modifications that enhance biological activity (i.e., epimerization at C20) also improve the direct contact of those analogues with the biologically important active site residues. Our results seem to lend support to the hypothesis that their interaction may be a requirement for the conformational changes leading to the formation of a highly stable heterodimer between VDR and RXR, as shown by digestion experiments.[15–17] Some of the LBD residues shown in Figure 1 may be among those that participate as "switches" that are activated upon contact with their ligand, inducing the structural changes that lead to an active VDR−RXR heterodimer. A detailed analysis of the single-residue interactions will be presented elsewhere. The robustness of the results and conclusions obtained from Table 2 will be checked by analyzing the single-residue energy distribution contributions to the contact energies displayed in Figure 2.

## References

(1) (a) Evans, R. M. The steroid and thyroid hormone receptor superfamily. *Science* **1988**, *240*, 889−895. (b) Kliewer, S. A.; Umesono, K.; Mangelsdorf, D. J.; Evans, R. M. Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling. *Nature* **1992**, *355*, 446−449.

(2) Darwish, H. M.; DeLuca, H. F. Recent advances in the molecular biology of vitamin D action. *Prog. Nucleic Acid Res. Mol. Biol.* **1996**, *53*, 321−344.

(3) Mangelsdorf, D. J.; Thummel, C.; Beato, M.; Herrlich, P.; Schütz, G.; Umesono, K.; Blumberg, B.; Kastner, P.; Mark, M.; Chambon, P.; Evans, R. M. The Nuclear Receptor Superfamily: The Second Decade. *Cell* **1995**, *83*, 835−839.

(4) Steinmetz, A. C.; Renaud, J. P.; Moras, D. Binding of ligands and activation of transcription by nuclear receptors. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 329−359.

(5) DeLuca, H. F.; Krisinger, J.; Darwish, H. The vitamin D system: 1990. *Kidney Int. Suppl.* **1990**, *29*, S2−S8.

(6) Reichel, H.; Koeffler, H. P.; Norman, A. W. The role of the vitamin D endocrine system in health and disease. *N. Engl. J. Med.* **1989**, *320*, 980−991.

(7) Bouillon, R.; Okamura, W. H.; Norman, A. W. Structure−Function Relationships in the Vitamin D Endocrine System. *Endocr. Rev.* **1995**, *16*, 200−257.

(8) Okamura, W. H.; Palenzuela, J. A.; Plumet, J.; Midland, M. M. Vitamin D: structure−function analyses and the design of analogs. *J. Cell. Biochem.* **1992**, *49*, 10−18.

(9) (a) Yamada, S.; Yamamoto, K.; Masuno, H.; Ohta, M. Conformation−function relationship of vitamin D: conformational analysis predicts potential side-chain structure. *J. Med. Chem.* **1998**, *41*, 1467−1475. (b) Zhou, X.; Zhu, G. D.; Van Haver, D.; Vandewalle, M.; De Clerq, P. J.; Verstuyf, A.; Bouillon, R. Synthesis, biological activity, and conformational analysis of four seco-D-15,19-bisnor-1α,25-dihydroxyvitamin D analogues diastereomeric at C17 and C20. *J. Med. Chem.* **1999**, *42*, 3539−3556.

(10) Rochel, N.; Wurtz, J. M.; Mitschler, A.; Klaholz, B.; Moras, D. The Crystal Structure of the Nuclear Receptor for Vitamin D Bound to Its Natural Ligand. *Mol. Cell* **2000**, *5*, 173−179.

(11) Tocchini-Valentini, G.; Rochel, N.; Wurtz, J. M.; Mitschler, A.; Moras, D. Crystal structures of the vitamin D receptor complexed to superagonist 20-epi ligands. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5491−5496.

(12) (a) Choi, M.; Yamamoto, K.; Masuno, H.; Nakashima, K.; Taga, T.; Yamada, S. Ligand Recognition by the Vitamin D Receptor. *Bioorg. Med. Chem.* **2001**, *9*, 1721−1730. (b) Choi, M.; Yamamoto, K.; Itoh, T.; Makishima, M.; Mangelsdorf, D. J.; Moras, D.; DeLuca, H. F.; Yamada, S. Interaction between Vitamin D Receptor and Vitamin D Ligands. Two-Dimensional Alanine Scanning Mutational Analysis. *Chem. Biol.* **2003**, *10*, 261−270.

(13) InsightII, Discover, and Search and Compare Biopolymer are trademarked software of Accelrys Inc., San Diego, CA.

(14) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: a program for macromolecular energy minimization and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.

(15) Liu, Y.-Y.; Collins, E. D.; Norman, A. W.; Peleg, S. Differential Interaction of 1α,25-Dihydroxyvitamin D3 Analogues and Their 20-*epi* Homologues with the Vitamin D Receptor. *J. Biol. Chem.* **1997**, *272*, 3336−3345.

(16) Peleg, S.; Sastry, M.; Collins, E. D.; Bishop, J. E.; Norman, A. W. Distinct conformational changes induced by 20-epi analogues of 1α,25-dihydroxyvitamin D3 are associated with enhanced activation of the vitamin D receptor. *J. Biol. Chem.* **1995**, *270*, 10551−10558.

(17) van den Bemd, G. C.; Pols, H. A.; Birkenhäger, J. C.; van Leeuwen, J. P. Conformational change and enhanced stabilization of the vitamin D receptor by the 1,25-dihydroxyvitamin D3 analog KH1060. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10685−10690.

(18) Ikekawa, N. Chemical synthesis of vitamin D analogues with selective biological activities. In *Vitamin D: Molecular, Cellular and Clinical Endocrinology*; Norman, A. W., Schaefer, K., Grigoleit, H. G., v Herrath, D., Eds.; Walter de Gruyter & Co.: Berlin, 1988; pp 25−33.

(19) Kubodera, N.; Watanabe, H.; Kawanishi, T.; Matsumoto, M. Synthetic studies of vitamin D analogues. XI. Synthesis and differentiation-inducing activity of 1α,25-dihydroxy-22-oxavitamin D3 analogues. *Chem. Pharm. Bull.* **1992**, *40*, 1494−1499.

(20) Binderup, L.; Latini, S.; Binderup, E.; Bretting, C.; Calverley, M.; Hansen, K. 20-Epi-vitamin D3 analogues: a novel class of potent regulators of cell growth and immune responses. *Biochem. Pharmacol.* **1991**, *42*, 1569−1575.

(21) Yang, W.; Freedman, L. P. 20-Epi analogues of 1,25-dihydroxyvitamin D3 are highly potent inducers of DRIP coactivator complex binding to the vitamin D3 receptor. *J. Biol. Chem.* **1999**, *274*, 16838−16845.

(22) Kubodera, N.; Miyamoto, K.; Akiyama, M.; Matsumoto, M.; Mori, T. Synthetic studies of vitamin D analogues. IX. Synthesis and differentiation-inducing activity of 1α,25-dihydroxy-23-oxa-, thia-, and azavitamin D3. *Chem. Pharm. Bull.* **1991**, *39*, 3221−3224.

(23) Murayama, E.; Miyamoto, K.; Kubodera, N.; Mori, T.; Matsunaga, I.; Synthetic studies of vitamin D3 analogues. VIII. Synthesis of 22-oxavitamin D3 analogues. *Chem. Pharm. Bull.* **1986**, *34*, 4410−4413.