Journal of Steroid Biochemistry & Molecular Biology 121 (2010) 34–42

Contents lists available at ScienceDirect

Journal of Steroid Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/jsbmb

MSITE: A new computational tool for comparison of homological proteins in holo form☆

Wanda Sicinska[a,*], Mateusz Kurcinski[b]

[a] Institute of Organic Chemistry, Polish Academy of Sciences, Kasprzaka 44/52, 01-224 Warsaw, Poland
[b] Department of Chemistry, University of Warsaw, 02-093 Warsaw, Poland

ARTICLE INFO

Article history:
Received 22 October 2009
Received in revised form 2 April 2010
Accepted 10 April 2010

Keywords:
Vitamin D receptor
Vitamin D
Vitamin D analogs
Structure comparisons
Targets for mutational analyses

ABSTRACT

The mechanism by which nuclear receptors respond differentially to structurally distinct agonists is not a well understood process. However, it is now obvious that transcriptional activity of nuclear receptors is a function of their interactions with co-activators. Recently, we released a new computational tool, CCOMP, for comparing side chain conformations in crystal structures of homologous protein complexes. Application of the CCOMP program revealed that 20-*epi*-1α,25-$(OH)_2D_3$ changes the side chain conformation of vitamin D receptor amino acids residing mostly far away from the ligand–receptor contacts. This strongly suggests that the ligand-co-activator signaling pathway involves indirect interactions between amino acids lining the binding pocket and outer surface residues that could attract co-activators. To facilitate identification of amino acids transmitting the subtle receptor changes upon ligand/modulator binding we developed another simple tool, MSITE. The program automatically lists the nearest neighbors of a given amino acid (for example neighbors of residues that are in contact with a ligand or reorient their side chains in the presence of a co-factor) in an arbitrary number of compared complexes. Comparison of seven binary vitamin D receptor complexes holding as ligands the analogs of 1α,25-$(OH)_2D_3$ with inverted configuration at carbon 14 or 20, or with incorporated oxolane ring bridging carbons 20 and 23, is reported.

1. Introduction

The vitamin D receptor (VDR) is a ligand-activated transcription factor sharing its 3D fold with other nuclear receptors (NRs) [1]. VDR's natural hormone, 1α,25-$(OH)_2D_3$ (1,25-D_3), regulates over 60 genes associated with calcium/phosphorus homeostasis, immune responses and cellular growth, differentiation or apoptosis [2]. Until now, only liganded VDR complexes have been crystallized [3]. The receptor's large ligand binding pocket (LBP) easily accommodates 1α,25-$(OH)_2D_3$ and its analogs with modified rings or side chains. It was found that irrespective of the structure, ligands anchor in the binding pocket in a similar fashion to 1α,25-$(OH)_2D_3$. Synthetic efforts to obtain potent vitamin D analogs mostly revolve around side chain modifications. The most potent agonists, KH1060 and MC1288, reveal transcription activity 200,000 and 100 times higher, respectively (Fig. 1 and Table 1S) than the natural ligand [4–14]. Superimposition of hVDR complexes holding as ligands KH1060 (1IE8) and MC1288 (1IE9) with 1,25-D_3-hVDR (1DB1) shows that the conformation of receptor backbones remains unchanged (Cα RMSD equal to 0.230 for 1IE8 and 0.243 for 1IE9), even when the vitamin D compounds differ so drastically in their transcriptional potency. These results suggest that co-activators might influence the conformation of the receptor's side chains (SC) and in this way modify VDR active sites responsible for transcription. Sometimes receptor–modulator interactions are critically dependent on just a few amino acids situated at the binding interface [15].

Although very informative, comparison of protein structures can be a time consuming process. Recently, we released a new computational tool, CCOMP [16,17], which compares the conformations of side chains in homologous protein complexes. Application of CCOMP [17] showed that 20-*epi*-1α,25-$(OH)_2D_3$ reorients most side chains of the VDR residues located far away from the ligand–receptor contacts. It is commonly accepted that mutations causing reduced binding to co-activators hamper transactivation. It is well documented that MC1288 interacts with the DRIP co-activator a 100 times more efficiently than 1α,25-$(OH)_2D_3$ and shows 500 times higher potency than the natural hormone in cell proliferation [2,5,6]. In order to provide theoretical proof that CCOMP has indeed identified surface residues whose change in conformation lead to a higher stability of the vitamin D receptor liganded by 20-*epi*-1α,25-$(OH)_2D_3$ (in comparison to the natural

☆ Special issue selected article from the 14th Vitamin D Workshop held at Brugge, Belgium on October 4–8, 2009.

* Corresponding author. Tel.: +48 22 8220211x310; fax: +48 22 6326681.
E-mail addresses: wbs@icho.edu.pl, wanda.sicinska@gmail.com (W. Sicinska).

0960-0760/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsbmb.2010.04.006

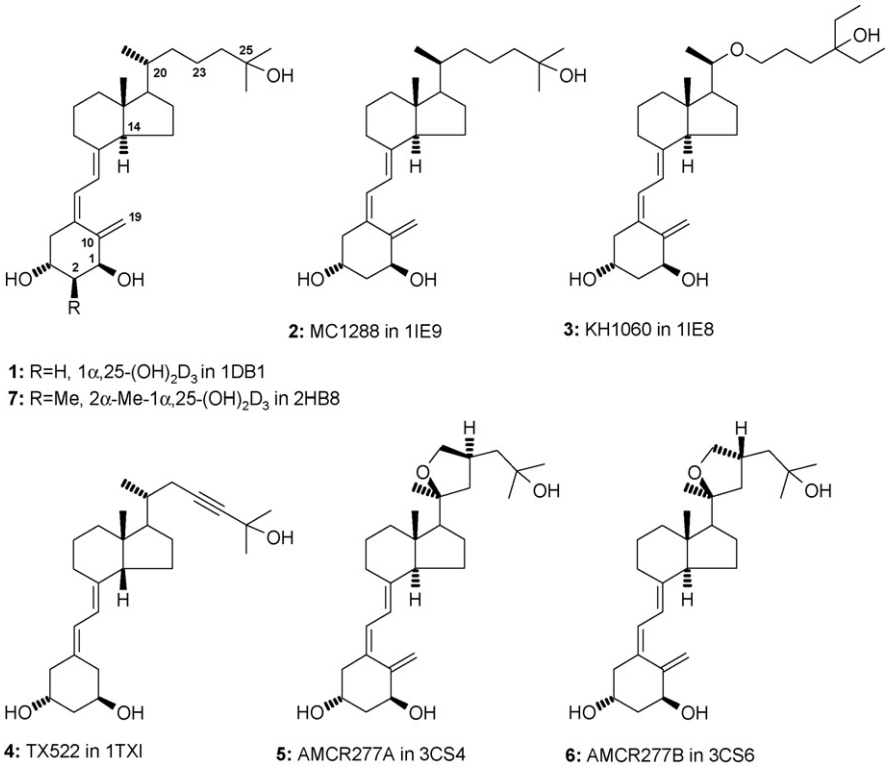


Fig. 1. Chemical structure of vitamin D ligands in studied complexes.

hormone), we performed peptide-protein docking experiments [17]. The fragment of SRC-1 co-activator containing the LxxLL motif was flexibly docked to the rigid VDR extracted from three complexes: 1,25-D_3-zVDRΔmt-SRC-1 (2HC4.pdb), 1,25-D_3-hVDRΔmt (1DB1.pdb) and MC1288-hVDRΔmt (1IE9.pdb). The fact that the energy of the SRC-1-$hVDR_{1IE9}$ complex (−343 kcal/mol) is lower than the energy of the SRC-1-$hVDR_{1DB1}$ complex (−337 kcal/mol) indicates that the receptor reoriented by 20-*epi*-1α,25-$(OH)_2D_3$ has a higher affinity to SRC-1 than VDR liganded by parental vitamin D.

Among the most important VDR's amino acids is hK264, which is conserved in NR and interacts with co-activators (CoA) from the p160 [18] and DRIP [19] families. In the vitamin D receptor K264 creates a salt bridge with E420, joining helices H4 and H12 [1]. This residue is exposed to solvent and possesses a different side chain conformation in 1DB1 and 1IE9 complexes, but a similar side chain conformation in 1IE9 and 2HC4 crystals. Therefore, it strongly suggests that MC1288 reorients the lysine's side chain preparing it for interactions with the SRC-1 co-activator. Thus, $K264_{1IE9}$ should interact with SRC-1 with less energetic cost than $K264_{1DB1}$. Taking into account that K264 is highly involved in the transcription process as is SRC-1 in up-regulation of osteocalcin (OC) gene [20], at least two functions (transcription and bone calcium mobilization) of the potent agonist 20-*epi*-1α,25-$(OH)_2D_3$ can be better understood.

There are some residues (for example D232, V418) in the vitamin D receptor which are important for transcription yet do not have their side chains reoriented in comparison with the 1,25-D_3-hVDR complex. These amino acids are hidden and therefore unable to directly interact with CoA (Tables 1 and 2). Their function is to keep the receptor in the transcriptionally active conformation [21–22]. While not detectable by CCOMP, these residues are part of hydrophobic chain interactions [1] and thus might trigger subtle receptor changes upon ligand binding. To help identify such amino acids, we developed another software tool, MSITE. The comparison of seven VDR complexes holding as ligands the analogs of 1α,25-$(OH)_2D_3$ (Fig. 1) that differ drastically in their biological functions (Table 1S) is reported here.

2. Computational methods

In this work we describe the results of pairwise comparisons of six holoVDR complexes (1IE9, 1IE8, 1TXI, 3CS4, 3CS6 and 2HB8) with 1DB1 serving as a reference (Fig. 1). In the reference complex, the vitamin D receptor holds as a ligand 1α,25-$(OH)_2D_3$. It is commonly accepted that biological activity of all vitamin D analogs are compared with this natural hormone.

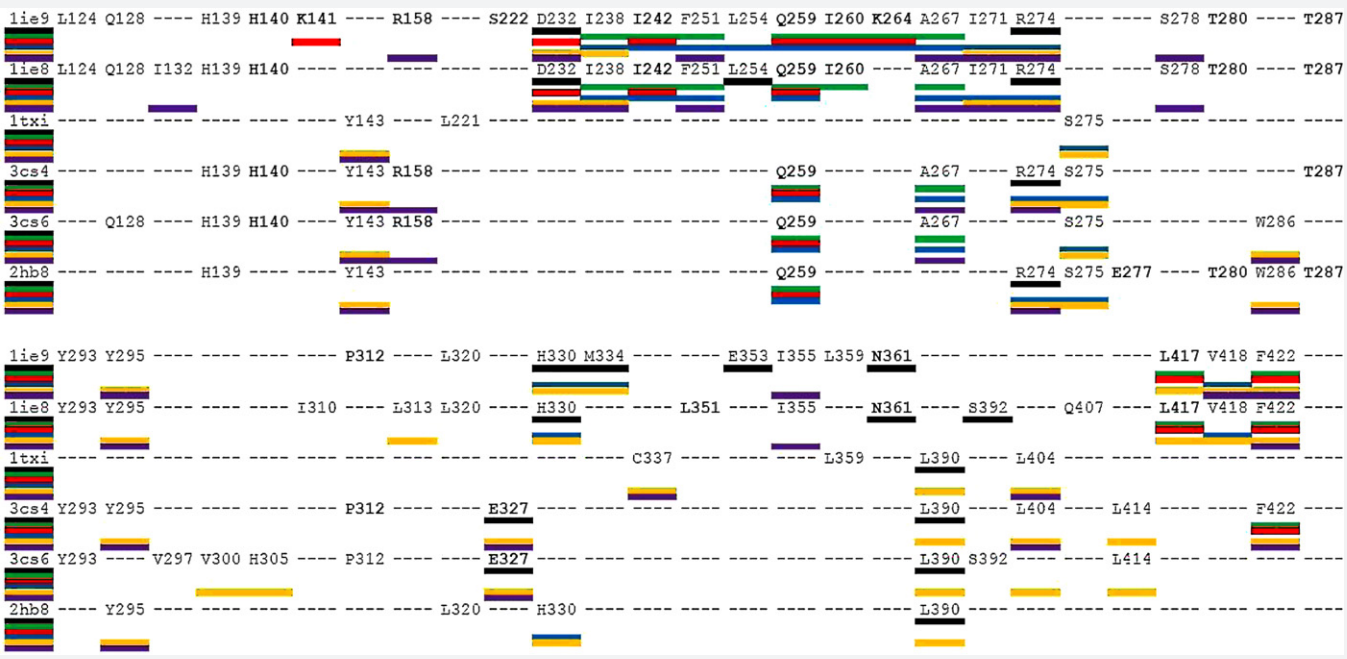
2.1. Computational approaches

Two recently developed computational tools were used for comparison of structural and biological properties of liganded VDR complexes: the CCOMP [17] and MSITE presented here. CCOMP finds the reoriented residues in VDR complexes holding different ligands, while MSITE detects changes in the nearest neighbors of selected amino acids.

2.1.1. CCOMP

CCOMP (Complex COMParison) performs several consecutive steps: aligning protein sequences, superimposing structures of the aligned proteins, transforming ligands according to the protein superposition, and calculating differences in orientation and conformation between individual amino acids and between the ligand molecules. While all of these steps can be done with the help of any interactive molecular visualization program, CCOMP automates the process and finally lists amino acids with significantly reoriented side chains, for which $\Delta\chi_{SC} \geq 10^{\circ}$. Additionally, the program can resolve data errors commonly encountered in Protein Data Bank (PDB) files, such as missing atoms or duplicated residues.

Table 1
VDR candidates for mutational experiments predicted by computational analysis: MSITE and CCOMP.



[a] All residues are numbered according to the hVDR sequence. Residues exposed to the surface with normalized solvent accessible surface above 20% are in bold style. Amino acids for which are available mutational experiments or known biological functions are underlined: in black residues sensitive to heterodimerization with RXR, in green and red to interactions with p160 and DRIP co-activators, in blue to transcription factors (TFs and TAFs) and in orange to transcription. Other amino acids important for VDR binding, holding of helical receptor structure or necessary for interaction with co-repressors are underlined in purple. Literature data are cited in Table 1S.

CCOMP calculates several difference measures: local all-atom RMSD (L-RMSD), amino acid center-of-mass deviation (C-RMSD) and distance RMSD per residue (D-RMSD), for whole amino acids and for side chains only. The L-RMSD reflects both spatial differences and conformational changes found in the compared complexes. The C-RMSD averages the atomic coordinates of particular amino acids, and therefore reflects mostly the spatial (translational) differences. On the other hand, the D-RMSD ignores translations and rotations and reflects only intra-residue conformational changes. To assess statistical significance of the difference measures, distributions of the RMSD values are normalized by their standard deviations. Yet another amino acid comparison measure implemented in CCOMP is the difference between values of side chain χ angles. These differences can reflect more subtle changes in side chain conformations than the D-RMSD values themselves. By default, all amino acids with χ angle differences exceeding 10° are listed by CCOMP. Taking into account the quality of the studied complexes (as quantified by B and R crystallographic factors), only the differences exceeding 40° are considered significant and discussed in this paper.

The results of CCOMP based comparison are summarized in Table 2. The hVDR complexes compared here differ in the biological activity of ligands. Only 19 amino acids of 250 receptor residues change their side chain orientation upon the replacement of 1α,25-$(OH)_2D_3$ with a vitamin D analog. Most of them (17) are located in the receptor's helices but none was found in AF2 (an activation function 2) helix (H12).

The main chain backbones of complexes compared in this paper superimpose with 1α,25-$(OH)_2D_3$-hVDR with an average accuracy of about 0.2 Å, with average values of Cα RMSD equal to 0.243 for 1IE9, 0.230 for 1IE8, 0.071 for 1TXI, 0.228 for 3CS4, 0.237 for 3CS6 and 0.189 for 2HB8, respectively. Values of local all-atom RMSD (L-RMSD) per residue, reflecting both spatial differences and conformational changes, are listed in Table 2. In the case of 'reoriented residues' L-RMSD values exceed 1.0.

2.1.2. MSITE

MSITE is a simple command-line program for immediate identification of the nearest neighbors of selected amino acids in an arbitrary number of compared complexes (Fig. 1S). The arguments of the MSITE program are the names of protein structures (in PDB format) and a list of residues to be analyzed. The residues are listed individually for each of the input PDB structures. The program outputs a list of amino acids in the neighborhood (distance between heavy atoms or protons ≤ 3.5 Å) of residues specified in the input list.

For MSITE's initial input, we used amino acids in contact with the ligand as well as residues that were identified by CCOMP as reoriented upon ligand binding. For each of these residues we compared neighborhoods in the 7 analyzed complexes. When an amino acid was not present in all neighborhoods we considered it as affected by ligand binding and appended it to the MSITE input list. This process was repeated until no new residues were identified. The results of our computational analysis are summarized in Tables 1 and 2.

2.2. Validation of crystal complexes and setting the cut off value for χ angle in CCOMP

Using MSITE we applied the same stringent criteria for the comparison process [17] as in the case of CCOMP. We studied the 1.5–2.0 Å resolution crystal structures of the liganded VDR complexes. They have an *R* factor of about 0.2 and belong to the same space group (Table 2S) [17]. It is reasonable to assume that crystal ordering and protein–surface interactions in complexes from the same space group would only differ due to changes induced by ligands. The Vitamin D receptor is very mobile in comparison with other nuclear receptors. The long flexible loop (residues 72–81 located between helices 1 and 3) hampers crystallization of the full-length LBD-VDR. For this reason all reported VDR X-ray structures are obtained for the liganded VDR deletion mutants. Even in high resolution crystallographic structure of the VDR,

Table 2
Residues predicted by CCOMP/MSITE as implicated in transcription activity and signaling mechanism in compared holoVDR complexes.[a]

Compared residues (motif)[b]		Compared complexes					
		1IE9	1IE8	1TXI	3CS4	3CS6	2HB8
L124	L-RMSD[c]	0.27	0.38				
(H1)	$\Delta\chi$ value[d]	8	3				
	SASA[e]	7	7				
Q128	L-RMSD[c]	1.08	1.01			1.78	
(H1)	$\Delta\chi$ value[d]	**176**	**166**			**167**	
	SASA[e]	19	16			17	
I132	L-RMSD[c]		0.91				
(H1)	$\Delta\chi$ value[d]		129				
	SASA[e]		5				
H139	L-RMSD[c]	1.15	1.15		1.15	1.15	1.15
(H1)	$\Delta\chi$ value[d]	**174**	**175**		**174**	173	**174**
	SASA[e]	3	1		1	1	1
H140	L-RMSD[c]	1.17	1.15		1.15	1.15	
(H1)	$\Delta\chi$ value[d]	**178**	**179**		**176**	**176**	
	SASA[e]	51	52		53	52	
K141	L-RMSD[c]	1.14					
(H1)	$\Delta\chi$ value[d]	**116**					
	SASA[e]	64					
Y143	L-RMSD[c]			0.07	0.12	0.12	0.11
(H1)	$\Delta\chi$ value[d]			2	1	1	1
	SASA[e]			6	5	6	5
R158	L-RMSD[c]	0.40			0.43	0.47	
($\Omega_{H2\text{-}H3n}$)	$\Delta\chi$ value[d]	**50**			**63**	**64**	
	SASA[e]	30			33	32	
L221	L-RMSD[c]			0.09			
(H3n)	$\Delta\chi$ value[d]			4			
	SASA[e]			16			
S222	L-RMSD[c]	0.85					
(H3n)	$\Delta\chi$ value[d]	**109**					
	SASA[e]	69					
D232	L-RMSD[c]	0.19	0.18				
(H3)	$\Delta\chi$ value[d]	3	2				
	SASA[e]	15	15				
I238	L-RMSD[c]	0.96	0.95				
(H3)	$\Delta\chi$ value[d]	**119**	**117**				
	SASA[e]	0	0				
I242	L-RMSD[c]	0.18	1.00				
(H3)	$\Delta\chi$ value[d]	2	**176**				
	SASA[e]	38	33				
F251	L-RMSD[c]	0.21	0.19				
(H4)	$\Delta\chi$ value[d]	1	3				
	SASA[e]	0	0				
L254	L-RMSD[c]	0.29	0.32				
(H4)	$\Delta\chi$ value[d]	5	4				
	SASA[e]	6	2				
Q259	L-RMSD[c]	1.07	1.09		1.08	1.08	1.07
(H4)	$\Delta\chi$ value[d]	**142**	**141**		**144**	**143**	**151**
	SASA[e]	20	19		16	17	18
I260	L-RMSD[c]	0.74	0.84				
(H4)	$\Delta\chi$ value[d]	**104**	**114**				
	SASA[e]	58	58				
K264	L-RMSD[c]	0.55					
(H4)	$\Delta\chi$ value[d]	**103**					
	SASA[e]	33					
A267	L-RMSD[c]	0.19	0.27		0.20	0.21	
(H5)	$\Delta\chi$ value[d,f]	NA	NA		NA	NA	
	SASA[e]	0	0		0	0	
I271	L-RMSD[c]	0.70	0.77				
(H5)	$\Delta\chi$ value[d]	**131**	**135**				
	SASA[e]	6	8				

Table 2 (*Continued*)

Compared residues (motif)[b]		Compared complexes					
		1IE9	1IE8	1TXI	3CS4	3CS6	2HB8
R274	L-RMSD[c]	0.08	0.06		0.08		0.11
(H5)	$\Delta\chi$ value[d]	2	2		3		1
	SASA[e]	7	5		6		6
S275	L-RMSD[c]			0.04	0.08	0.04	0.01
(H5)	$\Delta\chi$ value[d]			3	3	1	3
	SASA[e]			15	14	15	13
E277	L-RMSD[c]						0.14
(H5)	$\Delta\chi$ value[d]						8
	SASA[e]						37
S278	L-RMSD[c]	0.10	0.12				
(H5)	$\Delta\chi$ value[d]	0	2				
	SASA[e]	10	10				
T280	L-RMSD[c]	0.29	0.26				0.21
($\Omega_{H5\text{-}\beta}$)	$\Delta\chi$ value[d]	3	2				1
	SASA[e]	19	21				22
W286	L-RMSD[c]					0.18	0.16
(β1)	$\Delta\chi$ value[d]					2	1
	SASA[e]					6	6
T287	L-RMSD[c]	0.27	0.14		0.18		0.24
($\Omega_{H5\text{-}H6}$)	$\Delta\chi$ value[d]	8	3		1		3
	SASA[e]	46	42		41		41
Y293	L-RMSD[c]	0.46	0.37		0.63	0.59	
($\Omega_{H5\text{-}H6}$)	$\Delta\chi$ value[d]	32	24		**43**	36	
	SASA[e]	17	14		18	17	
Y295	L-RMSD[c]	0.32	0.41		0.27		0.11
($\Omega_{H5\text{-}H6}$)	$\Delta\chi$ value[d]	3	3		2		2
	SASA[e]	1	1		1		1
V297	L-RMSD[c]					1.372	
(H6)	$\Delta\chi$ value[d]					127	
	SASA[e]					42	
V300	L-RMSD[c]					0.39	
(H6)	$\Delta\chi$ value[d]					7	
	SASA[e]					11	
H305	L-RMSD[c]					0.15	
($\Omega_{H6\text{-}H7}$)	$\Delta\chi$ value[d]					8	
	SASA[e]					8	
I310	L-RMSD[c]		0.85				
(H7)	$\Delta\chi$ value[d]		139				
	SASA[e]		0				
P312	L-RMSD[c]	0.63			0.50	0.57	
(H7)	$\Delta\chi$ value[d]	**81**			**84**	**84**	
	SASA[e]	20			27	28	
L313	L-RMSD[c]		1.29				
(H7)	$\Delta\chi$ value[d]		143				
	SASA[e]		10				
L320	L-RMSD[c]	0.14	0.19				0.11
(H7)	$\Delta\chi$ value[d]	5	11				4
	SASA[e]	1	1				1
E327	L-RMSD[c]				0.19	0.18	
(H8)	$\Delta\chi$ value[d]				2	2	
	SASA[e]				35	34	
H330	L-RMSD[c]	0.13	0.11				0.07
(H8)	$\Delta\chi$ value[d]	7	3				3
	SASA[e]	7	8				8
M334	L-RMSD[c]	1.65					
(H8)	$\Delta\chi$ value[d]	**85**					
	SASA[e]	2					
C337	L-RMSD[c]			0.04			
(H8)	$\Delta\chi$ value[d]			1			
	SASA[e]			0			

Table 2 (*Continued*)

Compared residues (motif)[b]		Compared complexes					
		1IE9	1IE8	1TXI	3CS4	3CS6	2HB8
L351	L-RMSD[c]		1.28				
(H9)	$\Delta\chi$ value[d]		**140**				
	SASA[e]		42				
E353	L-RMSD[c]	0.31					
(H9)	$\Delta\chi$ value[d]	**43**					
	SASA[e]	51					
I355	L-RMSD[c]	0.83	0.94				
(H9)	$\Delta\chi$ value[d]	**105**	**116**				
	SASA[e]	7	5				
L359	L-RMSD[c]	0.19		0.05			
(H9)	$\Delta\chi$ value[d]	5		5			
	SASA[e]	0		0			
N361	L-RMSD[c]	0.32	0.23				
(H9)	$\Delta\chi$ value[d]	25	15				
	SASA[e]	59	59				
L390	L-RMSD[c]			0.49	1.29	1.29	0.53
(H10)	$\Delta\chi$ value[d]			**54**	**127**	**151**	**59**
	SASA[e]			0	0	0	0
S392	L-RMSD[c]		0.91			0.89	
(H11)	$\Delta\chi$ value[d]		120			115	
	SASA[e]		70			70	
L404	L-RMSD[c]			0.19	0.39	0.29	
(H11)	$\Delta\chi$ value[d]			9	7	4	
	SASA[e]			1	1	1	
L407	L-RMSD[c]		1.11				
(H11)	$\Delta\chi$ value[d]		172				
	SASA[e]		36				
L414	L-RMSD[c]				0.21	0.26	
($\Omega_{H11\text{-}H12}$)	$\Delta\chi$ value[d]				3	3	
	SASA[e]				4	4	
L417	L-RMSD[c]	0.31	0.293				
(H12)	$\Delta\chi$ value[d]	3	2				
	SASA[e]	26	32				
V418	L-RMSD[c]	0.27	0.24				
(H12)	$\Delta\chi$ value[d]	3	1				
	SASA[e]	3	3				
F422	L-RMSD[c]	0.35	0.19		0.24		
(H12)	$\Delta\chi$ value[d]	10	3		7		
	SASA[e]	11	12		13		

[a] All hVDR complexes are compared with 1α,25-$(OH)_2D_3$-VDR (1DB1) serving as a reference complex.
[b] Part of VDR to which the listed residue belongs is given in parenthesis. Helices (H) and loops (Ω) are numbered in accordance with notation used in Ref. [1].
[c] L-RMSD denotes local all-atom RMSD per whole amino acid.
[d] $\Delta\chi$ value denotes the difference between χ angles of side chains.
[e] SASA (given in %) denotes normalized solvent accessible surface area.

the conformations of flexible molecule parts have low reliability. The *B*-factor, also known as the "temperature factor," reflects the mobility of flexible substructures. High *B* values (exceeding 60) imply high uncertainty in a specific part of the model. Side chains are the most flexible elements in protein structures. For the creation of rotamer libraries only amino acids with *B*-factors below 40 are considered [23]. It is generally accepted that the position of a residuum is well defined if the *B*-factor of each atom of the studied protein fragment is below 40 [24]. The VDR amino acids compared in this paper fulfill these rather stringent criteria because only amino acids with a *B*-factor below 40 were considered (Table 3S).

It is accepted in the scientific community [25] that a *B*-factor of 20 corresponds to an error range of 0.5 Å in superimposed structures. *B*-factors of residue atoms exceeding 60 imply that side chain orientation is unknown. Taking these facts into account we set a cut off value of 40 for SC comparisons. Unfortunately, ternary VDR-SRC-1 and VDR-DRIP complexes have *B*-factor values exceeding 60 for almost all of their atoms, and therefore could not be reliably analyzed.

2.3. SASA (solvent accessible surface area)

The rearrangement of some VDR side chains upon ligand binding prepares protein complexes for gene transcription regulated by this receptor. Since the transcription process is managed by co-activators, knowledge of whether the reoriented residues are exposed to a solvent and are capable of attracting CoA is of significant interest. In this work values of solvent accessible surface area (SASA) were calculated using the DSSP program [26] and then normalized by the average amino acid surface [17]. The average surface value per amino acid is defined as the surface of the central residue in reference AXA tripeptides [27]. Table 4S contains normalized SASA values of hVDR residues in the studied complexes. For all amino acids listed in Table 1 we also placed their SASA values in Table 2.

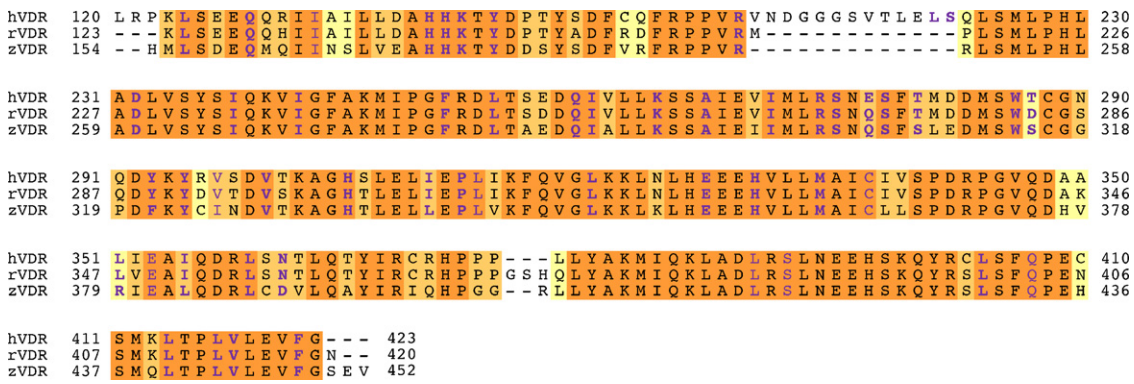


Fig. 2. Alignments of human, rat and zebrafish VDR sequences (LBD fragment) generated by PSI-BLAST. The orange, dark yellow and yellow bars highlight, respectively residues with high, moderate and low conservancy. The fifty-three residues predicted by MSITE and CCOMP programs as candidates for mutational experiments are marked in purple. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Twenty of the 53 residues found by MSITE/CCOMP are exposed to solvent and are therefore capable of attracting a co-modulator protein or RXR heterodimerization partner. Most of them (17) occupy helices: H1 (H140 and K141), H3n (S222), H3 (I242), H4 (Q259, I260 and K264), H5 (E277), H6 (V297), H7 (P312), H8 (L351 and E327), H9 (E351, E353 and N361), H11 (S392 and L407), H12 (L417). Mutational experiment data are available for six out of the seventeen residues. In all six cases the mutated amino acids impaired the VDR affinity to co-activators.

As of yet, VDR–RXR complexes have not been crystallized. Fortunately, the crystallographic structure of the RXR heterodimers with other nuclear receptors has been analyzed [28,29]. It has been observed that structural elements forming the dimerization interfaces in the NR family are identical [30] and contain residues from helices H7, H9, H10 and H11, as well as residues from loops 8-9 and 9-10. The core of the heterodimerization interface contains amino acids from two helices: H9 and H10. There is a body of evidence that the VDR–RXR interface is similar to that observed in the majority of the nuclear receptors [30]. Heterodimerization experiments performed with wild type [31] and mutated VDR support this hypothesis [21,32]. By analogy to RAR-RXR and PPAR-RXR complexes we can expect that VDR–RXR dimerization interface may involve several vitamin D receptor residues: I336, P341, D342, P344, E353, Q356, D357, S360, N361, Q364, R368, Q378, Y387, A381, K382, I384, Q385, K386, L387, A388, D389, R391, S392 and E395. Three of them were found by MSITE/CCOMP as reoriented (E353 and S392) or having a different environment (N361) in the transcriptionally very potent complexes 1IE9 and/or 1IE8.

2.4. Multiple sequence alignment

The sequences of human, rat and zebrafish VDRs were aligned using the PSI-BLAST procedure [33]. The alignments covering residues 120–423 of the hVDR are shown in Fig. 2. Table 2S lists PDB codes of the proteins used for alignment. Only two (T287 and L351) of the fifty-three residues predicted by CCOMP/MSITE as candidates for mutational experiments are not conserved in a VDR family.

3. Results and discussion

The vitamin D receptor acts as a ligand-dependent transcription factor. Transcription of genes controlled by nuclear receptors is a multistep process that is still poorly understood. It is speculated that the binding of VDR to RXR increases its affinity to the vitamin D receptor response element (VDRRE) [34]. When a ligand–VDR–RXR complex reaches cognate genes, sequential/combinatorial actions of co-activators start. First to act are the p160 modulators, followed by the mediator-DRIP and finally DNA splicing modulators TIFs and TAFs are attracted [35–37]. It is well documented that VDR mutations that reduce binding to co-modulators also hamper transactivation [18,30,36]. So far, residues from three VDR helices: H10 [36], H11 [38], and H12 [31], as well as from the groove between H3 and H5, were identified as the main co-activator interfaces. In the past decade several complexes of the human, rat and zebrafish VDR have been successfully crystallized in the absence and presence of peptides mimicking the short sequence of SRC-1 and DRIP co-activators [1,19,39,40]. Superimposition of these complexes demonstrated that conformations of receptor backbones and the topology of ligands in the ligand binding pocket (LBP) of the VDR are preserved, even if the vitamin D compounds have different transcription abilities [3].

To find subtle changes in the receptor's structure that results from ligand/co-activator binding it is extremely helpful to employ a computer aided complex comparison. Recently, we released a new computational tool, CCOMP [17], which lists as output amino acids having reoriented side chains in the proteins compared. The application of CCOMP to VDR liganded with 20-*epi*-1α,25-$(OH)_2D_3$ revealed conformational changes in amino acids residing mostly far away from the ligand–receptor interface. This strongly suggests that the ligand-CoA signaling pathway involves indirect interactions between amino acids lining the binding pocket and outer surface residues capable of attracting co-activators. To facilitate identification of amino acids that could be part of the signaling pathway, we developed another simple software tool, MSITE. The program outputs the nearest neighbors of amino acids listed in the input file. In this paper we track differences in the neighborhoods of VDR residues liganded by 1α,25-$(OH)_2D_3$ and its six analogs modified on carbon 14, 20 and 23 (Fig. 1). Such structural changes (inversion of configuration at C14 or C20 or incorporation of oxolane ring bridging C20 and C23) drastically alter the biological profile of vitamin D compounds (Table 1S).

The comparison of six vitamin D receptor complexes (Table 2) revealed that liganding affects 53 out of 250 LBP-VDR residues (either through conformation or neighborhood changes). Only two (T287 and L351) of them are not likely to be functional since they are not conserved in VDR family (Fig. 2). Residues most "responsive" to ligand changes occupy helices 4, 5 and 12. These helices have been identified as forming co-factor binding surfaces in the family of nuclear receptors [15]. Almost half of the residues affected by ligands have differently oriented side chains (24) and/or are not surrounded by the same amino acids (29) in comparison with 1,25-D_3-VDR. Thirty-four out of the 53 residues are known to be important in the transcription process (Tables 1 and 2). They are involved in various transcription events; 10 residues (D232, L254, R274, E327, H330 M334, E353, N361, L390 and S392) participate in heterodimerization with RXR [28,29], 18 residues (K141, D232,

I238, I242, F251, L254, Q259, I260, K264, A267, I271, R274, S275, H330, M334, L417, V418 and F422) are capable of interacting with co-activators from p160/DRIP/TAF/TIF families [12,17,19–21,41], 9 residues (Y143, I238, K264, I271, R274, W286, Y295, L404 and F422) are essential for holding the VDR in active agonistic conformation [22] and 6 residues (F251, K264, A267, Y297, L390 and L417) participate in allosteric communication between four functional LBP-NR surfaces [15]. This demonstrates that CCOMP/MSITE is highly successful in finding functional amino acids. Given that KH1060 is a superagonist and MC1288 is a strong agonist in relation to $1,25(OH)_2D_3$ (Table 1S), is reasonable to expect that their complexes (1IE8 and 1IE9) should show the largest number of differences relative to 1DB1. This is indeed the case (Table 2) since 1IE8, 1IE9, 1TXI, 3CS4, 3CS6, and 2HB8 complexes reveal differences in relation to the reference complex 1DB1 on 32, 33, 7, 16, 19, and 13 amino acids, respectively. In the case of the superagonist's complex, three residues (I310, L313, and Q407) are specific, i.e. emerge only once in the six compared complexes. It is possible that these amino acids are responsible for the exceptional increase in the transcriptional potency of KH1060 and thus present good targets for structure-function studies. Selective differences were found in 1IE9 on 5 residues: K141, S222, K264, M334 and E353; all have reoriented side chains in comparison to 1DB1 and all except for M334 are exposed to solvent. Since 2 of 5 residues (K141 and K264) are known to interact with p160 and/or DRIP proteins, it is likely that they are responsible for enhancing the cellular differentiation and proliferation of the MC1288 compound.

By jointly applying CCOMP/MSITE we found differences on 27 amino acids (L124, G129, I132, H139, K141, F150, R158, S222, I242, S256, K264, Y293, Y295, I310, P312, K315, K322, N324, M334, C337, L351, Q364, K386, S392, E395, C403 and Q407) while comparing the VDR complexes holding the superagonist KH1060 and the strong agonist MC1288. Sparse biological data known for these complexes (Table 1S) indicate that they differ primarily in transcriptional potency. We believe that some of the 27 residues are instrumental for the tremendous transcriptional activity of KH1060. Of the 27 residues, 17 are exposed to solvent (SASA ≥ 20%), they are depicted in Fig. 3.

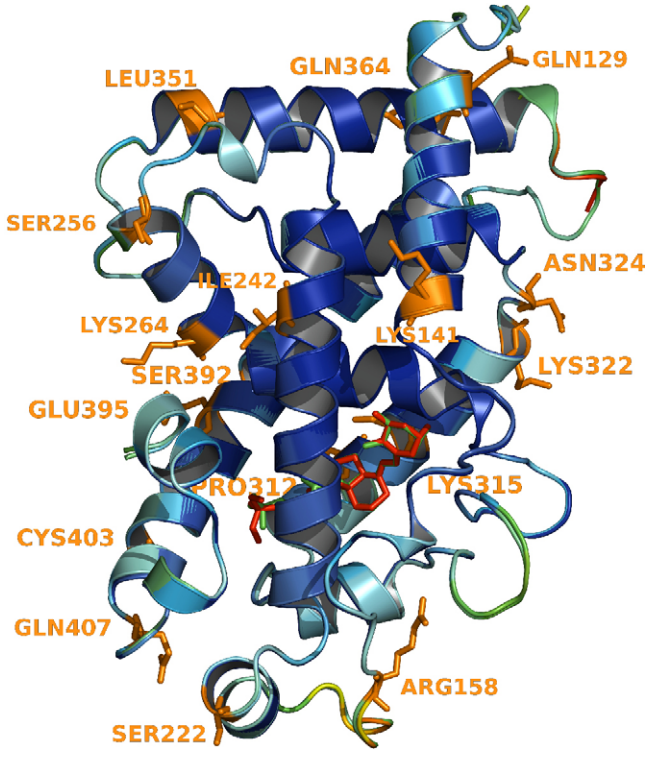


Fig. 3. Superimposition of two VDR complexes: MC1288-hVDRmt (1IE9, ligand in green) and KH1060-hVDRmt (1IE8, ligand in red). Helices are colored according to *B*-factor of amino acids; rigid residues with small *B* values are depicted in dark blue. Seventeen residues are marked, as indicated by CCOMP/MSITE output, possessing different side chain's orientation or neighbors. Residues contacting co-factor or reoriented after changing the vitamin D ligand (MC1288 versus KH1060) were used as input to MSITE. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Limited mutational data for the studied complexes precludes us from identifying specific co-activators responsible for each transcription step. However, available experimental data strongly suggest that amino acids found by the CCOMP/MSITE are important for transcription events (Tables 1 and 2). About 15 amino acids identified by CCOMP/MSITE have not been previously reported as playing a role in the activation of the VDR by ligands. We believe that the residues for which no biological data exists might also be important and thus constitute good targets for mutagenesis. Comprehensive analysis of VDR-CoA interactions performed on receptors mutated at positions proposed in this study could greatly enrich our knowledge of protein fragments responsible for functional specificity.

4. Conclusions

Although very informative, comprehensive comparison of protein structures can be a cumbersome process. The computational tools we developed (CCOMP and MSITE) automate two processes: the comparison of protein complexes and identification of target amino acids for genetic engineering. These programs are suitable for comparison of any liganded/unliganded homological protein complexes.

Detailed comparison of human VDR complexes holding as ligands vitamin D compounds differing in their biological potency produced a relatively small number (around 20%) of residues affected by ligand changes. The majority of residues are involved in processes leading to transcription of genes controlled by the vitamin D receptor. Some experimental data are available for 34 out of the 53 amino acids listed in Table 1 and all 34 are involved in transcription events. This strongly suggests that residues found by CCOMP and MSITE are functionally important rather than randomly chosen due to measurement error. We believe that both programs can be very useful in searching for structure-function relationships in proteins.

Acknowledgments

W.S. would like to express her thanks to Professor Andrzej Kolinski for critical reading of this manuscript and Dr. Piotr Rotkiewicz for participation in developing of MSITE program.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jsbmb.2010.04.006.

References

[1] N. Rochel, J.M. Wurtz, A. Mitschler, B. Klaholz, D. Moras, The crystal structure of the nuclear receptor for vitamin D bound to its natural ligand, Mol. Cell 5 (2000) 173–179.
[2] S. Yamada, M. Shimizu, K. Yamamoto, Structure–function relationships of vitamin D including ligand recognition by the vitamin D receptor, Med. Res. Rev. 23 (2003) 89–115.
[3] N. Rochel, D. Moras, Ligand binding domain of vitamin D receptors, Curr. Top. Med. Chem. 6 (2006) 1229–1241.
[4] R. Bouillon, W.H. Okamura, A.W. Norman, Structure–function relationships in the vitamin D endocrine system, Endocr. Rev. 16 (1995) 200–257.

[5] H.F. DeLuca, J. Wicha, 20(S)-25-hydroxy vitamin D compounds, US Patent 5,840,938 (1998).
[6] W. Yang, L.P. Freedman, 20-*epi* analogues of 1,25-dihydroxyvitamin D_3 are highly potent inducers of DRIP coactivator complex binding to the vitamin D_3 receptor, J. Biol. Chem. 274 (1999) 16838–16845.
[7] F.A. Weyts, P. Dhawan, X. Zhang, J.E. Bishop, M.R. Uskokovic, Y. Ji, G.P. Studzinski, A.W. Norman, S. Christakos, Novel Gemini analogs of 1α,25-dihydroxyvitamin D_3 with enhanced transcriptional activity, Biochem. Pharmacol. 67 (2004) 1327–1336.
[8] S. Peleg, M. Sastry, E.D. Collins, J.E. Bishop, A.W. Norman, Distinct conformational changes induced by 20-*epi* analogues of 1α,25dihydroxyvitamin D_3 are associated with enhanced activation of the vitamin D receptor, J. Biol. Chem. 270 (1995) 10551–10558.
[9] L. Binderup, E. Bretting, C. Calverley, M. Hansen, 20-epi-vitamin D analogues: a novel class of potent regulators of cell growth and immune responses, Biochem. Pharmacol. 42 (1991) 1569–1575.
[10] N. Tocchini-Valentini, J.M. Rochel, D. Wurtz, Moras, Crystal structures of the vitamin D nuclear receptor liganded with the vitamin D side chain analogues calcipotriol and seocalcitol, receptor agonists of clinical importance, Insight into a structural basis for the switching of calcipotriol to a receptor antagonist by further side chain modification, J. Med. Chem. 47 (2004) 1956–1961.
[11] G. Eelen, L. Verlinden, M.V. Camp, F. Claessens, P. De Clercq, M. Vandewalle, R. Bouillon, A. Verstuyf, Altered vitamin D receptor–coactivator interactions reflect superagonism of vitamin D analogs, J. Steroid Biochem. Mol. Biol. 97 (2005) 65–68.
[12] G. Eelen, L. Verlinden, N. Rochel, F. Claessens, P. De Clercq, M. Vandewalle, G. Tocchini-Valentini, D. Moras, R. Bouillon, A. Verstuyf, Superagonistic action of 14-epi-analogs of 1,25-dihydroxyvitamin D explained by vitamin D receptor coactivator interaction, Mol. Pharmacol. 67 (2005) 1566–1573.
[13] S. Hourai, L.C. Rodrigues, P. Antony, B. Reina-San-Martin, F. Ciesielski, B.C. Magnier, K. Schoonjans, A. Mourino, N. Rochel, D. Moras, Structure-based design of a superagonist ligand for the vitamin D nuclear receptor, Chem. Biol. 15 (2008) 383–392.
[14] S. Hourai, T. Fujishima, A. Kittaka, Y. Suhara, H. Takayama, N. Rochel, D. Moras, Probing a water channel near the A-ring of receptor-bound 1α,25-dihydroxyvitamin D_3 with selected 2α-substituted analogues, J. Med. Chem. 49 (2006) 5199–5205.
[15] A.I. Shulman, C. Larson, D.J. Mangelsdorf, R. Ranganathan, Structural determinants of allosteric ligand activation in RXR heterodimers, Cell 116 (2004) 417–429.
[16] CCOMP program, http://www.icho.edu.pl/Eng/Achievements.
[17] W. Sicinska, P. Rotkiewicz, Structural changes of vitamin D receptor induced by 20-*epi*-1α,25-$(OH)_2D_3$: an insight from a computational analysis, J. Steroid Biochem. Mol. Biol. 113 (2009) 253–258.
[18] W. Feng, R.C.J. Ribeiro, R.L. Wagner, H. Nguyen, J.W. Apriletti, R.J. Fletteric, J.D. Baxter, P.J. Kushner, B.L. West, Hormone-dependent co-activator binding to a hydrophobic cleft on nuclear receptors, Science 280 (1998) 1747–1749.
[19] J.L. Vanhooke, M.M. Benning, B.C. Bauer, J.W. Pike, H.F. DeLuca, Molecular structure of the rat vitamin D receptor ligand binding domain complexed with 2-carbon-substituted vitamin D_3 hormone analogues and a LXXLL-containing co-activator peptide, Biochemistry 43 (2004) 4101–4110.
[20] L. Carvallo, B. Henriquez, J. Olate, A.J. van Wijnen, J.B. Lian, G.S. Stein, S. Onate, J.L. Stein, M. Montecino, The 1α,25-dihydroxy vitamin D_3 receptor preferentially recruits the coactivator SRC-1 during up-regulation of the osteocalcin gene, J. Steroid Biochem. Mol. Biol. 103 (2007) 420–424.
[21] S. Vaisanen, S. Ryhanen, J.T.A. Saarela, M. Perakyla, T. Andresin, P.H. Maenpaa, Structurally and functionally important amino acids of the agonistic conformation of the human vitamin D receptor, Mol. Pharmacol. 62 (2004) 788–794.
[22] S. Yamada, K. Yamamoto, Ligand recognition by vitamin D receptor: total alanine scanning mutational analysis of the residues lining binding pocket of vitamin D receptor, Curr. Top. Med. Chem. 6 (2006) 1255–1265.
[23] R.L. Dunbrack, Rotamer libraries in the 21st century, Curr. Opin. Struct. Biol. 12 (2002) 431–440.
[24] S.C. Lovell, J.M. Word, J.S. Richardson, D.C. Richardson, Asparagine and glutamine rotamers: *B*-factor cutoff and correction of amide flips yield distinct clustering, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 400–405.
[25] http://www.usm.maine.edu/∼rhodes/SPVTut/text/STut09aTN.html.
[26] W. Kabsh, C. Sander, DSSP, updated CMBI version by ElmK, 1982.
[27] C. Chothia, The nature of the accessible and buried surfaces in proteins, J. Mol. Biol. 105 (1976) 1–14.
[28] W. Bourguet, V. Vivat, J.M. Wurtz, P. Chambon, H. Gronemeyer, D. Moras, Crystal structure of a heterodimeric complex of RAR and RXR ligand-binding domains, Mol. Cell 5 (2000) 289–298.
[29] R.T. Gampe, V.G. Montana, M.H. Lambert, A.B. Miller, R.K. Bledsoe, M.V. Milburn, S.A. Kliewer, T.M. Wilson, H.E. Xu, Asymmetry in the PPARγ/RXRα crystal structure reveals the molecular basis of heterodimerization among nuclear receptors, Mol. Cell 5 (2000) 545–555.
[30] P. Germain, B. Staels, C. Dacquet, M. Spedding, V. Laudet, International Union of Pharmacology. LXIII. Retinoid X Receptors, Pharmacol. Rev. 58 (2006) 685–704.
[31] P. Pathrose, O. Barmina, C.Y. Chang, D.P. MacDonell, N.K. Shevde, W. Pike, Inhibition of 1,25-dihydroxyvitamin D_3-dependent transcription by synthetic LXXLL peptide antagonists that target the activation domains of the vitamin D and retinoid X receptors, J. Bone Miner. Res. 17 (2002) 2196–2205.
[32] C.H. Jin, S.A. Kerner, M.H. Hong, J.W. Pike, Transcriptional activation and dimerization functions in the human vitamin D receptor, Mol. Endocrinol. 10 (1996) 945–957.
[33] S.F. Atschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[34] R. Sanchez-Martinez, A.I. Castillo, A. Steinmeyer, A. Aranda, The retinoid X receptor ligand restores defective signaling by the vitamin D receptor, EMBO Rep. 7 (2006) 1030–1034.
[35] C. Rachez, L.P. Freedman, Mechanisms of gene regulation by vitamin D_3 receptor: a network of co-activator interactions, Gene 246 (2000) 9–21.
[36] J.B. Barry, G.M. Leongs, W.B. Church, L.L. Issa, J.A. Eisman, Interactions of SKIP/NCoA-62, TFIIB, and Retinoid X Receptor with vitamin D receptor helix H10 residues, J. Biol. Chem. 278 (2003) 8224–8228.
[37] B.D. Darimont, R.L. Wagner, J.W. Apriletti, M.R. Stallcup, P.J. Kushner, J.D. Baxter, R.J. Fletterick, K.R. Yamamoto, Structure and specificity of nuclear receptor–coactivator interactions, Genes Dev. 12 (1998) 3343–3356.
[38] W. Lee, N. Noy, Interaction of RXR with co-activators are differentially mediated by helix 11 of the receptor's ligand binding domain, Biochemistry 41 (2004) 2500–2508.
[39] G.D. Tocchini-Valentini, N. Rochel, J.M. Wurtz, A. Mitschler, D. Moras, Crystal structures of the vitamin D receptor complexed to superagonist 20-epi ligands, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 5491–5496.
[40] F. Ciesielski, N. Rochel, D. Moras, Adaptability of the vitamin D nuclear receptor to the synthetic ligand Gemini: remodeling the LBP with one side chain rotation, J. Steroid Biochem. Mol. Biol. 103 (2007) 235–242.
[41] M. Herdick, Y. Bury, M. Quack, M.R. Uskokovic, P. Polly, C. Carlberg, Response element and coactivator-mediated conformational change of the vitamin D_3 receptor permits sensitive interaction with agonists, Mol. Pharmacol. 57 (2000) 1206–1217.