Journal of Molecular Graphics and Modelling 28 (2009) 287–296

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

# Investigating the differential activation of vascular endothelial growth factor (VEGF) receptors

Bruno Araújo Cautiero Horta [a,*], Ana Carolina Rennó Sodero [b], Ricardo Bicca de Alencastro [a,**]

[a] Physical Organic Chemistry Group, Departamento de Química Orgânica, Instituto de Química, Universidade Federal do Rio de Janeiro, Cidade Universitária, CT, Bloco A, lab. 609, Rio de Janeiro, RJ 21941-909, Brazil

[b] Departamento de Bioquímica e Biologia Molecular, Fundação Oswaldo Cruz, Instituto Oswaldo Cruz, Av. Brasil, 4365. Pavilhão Leonidas Deanne, Laboratório de Bioquímica de Proteínas e Peptídeos, Sl. 309, Manguinhos, Rio de Janeiro, RJ 2104590, Brazil

ARTICLE INFO

*Article history:*
Received 6 March 2009
Received in revised form 9 August 2009
Accepted 10 August 2009
Available online 15 August 2009

*Keywords:*
VEGF
Angiogenesis
Molecular dynamics
Receptor-binding
Homology modeling
Surface analysis
Hydrogen bond

ABSTRACT

The vascular endothelial growth factors are key mediators of angiogenesis and are also related to several physiological processes such as monocyte chemotaxis, dendritic cell development, hematopoietic stem cell survival, and many others. PlGF, VEGF, VEGFB, VEGFC and VEGFD were identified as members of the vascular endothelial growth factor family. They act by differential activation of three receptors: Flt-1, KDR and Flt-4. PlGF and VEGFB only activate Flt-1. VEGF activates both Flt-1 and KDR. VEGFC and VEGFD activate KDR and Flt-4. The available three dimensional structures of VEGF and PlGF, in complex with the domain-2 of Flt-1, show that both proteins bind in a very similar way to Flt-1 receptor. Here we construct the three dimensional model of the domain-2 of KDR receptor using the same domain of Flt-1 as template. We also construct the model complexes VEGF/KDR, VEGFB/Flt-1, VEGFB/KDR and PlGF/KDR. Molecular dynamics simulations with explicit solvent are carried out on eleven molecular systems: unbound VEGF, VEGF/Flt-$1_{D2}$, VEGF/$KDR_{D2}$, unbound PlGF, PlGF/Flt-$1_{D2}$, PlGF/$KDR_{D2}$, unbound VEGFB, VEGFB/Flt-$1_{D2}$, VEGFB/$KDR_{D2}$, unbound Flt-$1_{D2}$ and unbound $KDR_{D2}$. We analyze protein–protein interactions, shape complementarity, charge complementarity and hydrogen bonds. As a coarse estimation of the desolvation penalties, we assume a correlation to the number of hydrogen bonds with solvent molecules that are lost upon complex formation. The results herein are consistent with the experimental selectivity profile (VEGF being able to activate both Flt-1 and KDR receptors while VEGFB and PlGF being only able to activate Flt-1), and provide a collection of evidences sustaining the complementarity of polar interactions as the main responsible for protein recognition and selectivity.


## 1. Introduction

The vascular endothelial growth factors (VEGFs) are key mediators of angiogenesis and are also related to several physiological processes such as monocyte chemotaxis, dendritic cell development, hematopoietic stem cell survival, among others [1,2]. Several pathological phenomena are also caused or aggravated by the expression of VEGFs [3,4]. VEGFs are polypeptides that belong to the cystine-knot superfamily of signaling proteins, which, in humans, include placental growth factor (PlGF), vascular endothelial growth factor-A (VEGFA, which is commonly known simply as vascular endothelial growth factor or VEGF), vascular endothelial growth factor-B (VEGFB), vascular endothelial growth factor-C (VEGFC) and vascular endothelial growth factor-D (VEGFD) [5,6]. They act by differential activation of three transmembrane tyrosine kinase receptors: Flt-1 (or VEGFR-1), KDR (or VEGFR-2) and Flt-4 (or VEGFR-3) [7]. Both KDR and Flt-1 receptors are essential to embryonic development. In adults, KDR seems to be the most important vascular endothelial growth factor receptor and its activation promotes endothelial cell migration, *in vivo* angiogenesis and increased vascular permeability. Activation of the Flt-1 receptor induces monocyte and macrophage migration and the sequestering of excess VEGF [1]. Activation of the Flt-4 receptor is mainly related to the growth of lymphatic vessels, which is known to be associated with the metastatic spread of tumor cells [8]. The complex biological responses are promoted by the activation of one or more receptors. This fine selection is a consequence of the molecular complementarity between the different factors and receptors. VEGF is able to activate both Flt-1 and KDR receptors. PlGF and VEGFB are only able to activate Flt-1 receptor. Mature form of VEGFC and VEGFD can activate both KDR and Flt-4 [1,8,9].

* Corresponding author. Current address: Laboratory of Physical Chemistry, Eidgenössische Technische Hochschule, Hönggerberg, HCI, CH-8093 Zürich, Switzerland. Tel.: +55 21 2562 7132; fax: +55 21 2562 7132.
** Corresponding author. Tel.: +55 21 2562 7132; fax: +55 21 2562 7132.
*E-mail addresses:* bruno.horta@igc.phys.chem.ethz.ch (B.A.C. Horta), bicca@iq.ufrj.br (R.B.d. Alencastro).

1093-3263/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2009.08.007

The 3-dimensional structures of VEGF (PDB codes: 1VPF [2.5 Å resolution] and 2VPF [1.93 Å resolution]) [6,10], PlGF (PDB code: 1FZV [2.0 Å resolution]) [11] and VEGFB (PDB code: 2C7W [2.48 Å resolution]) [12] were previously determined and are available at the Protein Data Bank. Wiesmann and co-workers showed that the second and third domains of Flt-1 are necessary and sufficient for binding to VEGF and that the domain-2 alone binds only 60-fold less tightly than the full receptor [13]. The 3-dimensional structures of VEGF complexed with the domain-2 of the Flt-1 receptor (Flt-$1_{D2}$) showed that two copies of Flt-$1_{D2}$ bind to VEGF (PDB code: 1FLT [1.7 Å resolution] and 1QTY [2.7 Å resolution]) [13,14]. Similarly, the determination of a crystal structure showed that PlGF also binds to two copies of Flt-$1_{D2}$ with practically the same orientation adopted by VEGF (PDB code: 1RV6 [2.45 Å resolution]) [15]. The determination of these crystal structures provided valuable information on how VEGF and PlGF interact with Flt-1 receptor. However, the structure of the same domain of KDR receptor is as yet unknown and, consequently, how VEGF interacts with it still remains unclear. We expect the orientation of the proteins in the VEGF/KDR complex to be very similar to that observed for the VEGF/Flt-1 complex (i.e. there is a 40% sequence identity between the domain-2 of the two receptors, which probably indicates a very similar 3-dimensional structure). However, it is known that PlGF and VEGFB are unable to activate KDR (i.e. probably the complexes PlGF/KDR and VEGFB/KDR are not stable enough to promote activation or for some reason they do not form at all). The elucidation of the 3-dimensional structure of KDR and the study of the interactions between KDR and growth factors could contribute to the understanding of the selectivity profile.

In the present work, we generate a 3-dimensional structural model of the second domain of KDR, $KDR_{D2}$, based on sequence homology to the same domain of Flt-1. Models of the complexes VEGF/$KDR_{D2}$, VEGFB/Flt-$1_{D2}$, VEGFB/$KDR_{D2}$ and PlGF/$KDR_{D2}$ are constructed by superimposing the structures of individual proteins onto the known crystal structures of complexes VEGF/Flt-$1_{D2}$ and PlGF/Flt-$1_{D2}$. Molecular dynamics (MD) simulations are carried out for all of these systems. A total of eleven MD trajectories are analyzed: unbound factors VEGF, VEGFB and PlGF; unbound receptors Flt-$1_{D2}$ and $KDR_{D2}$; and the complexes VEGF/Flt-$1_{D2}$, VEGF/$KDR_{D2}$, VEGFB/Flt-$1_{D2}$, VEGFB/$KDR_{D2}$, PlGF/Flt-$1_{D2}$ and PlGF/$KDR_{D2}$. We investigate why VEGF is able to activate both Flt-1 and KDR receptors while PlGF and VEGFB are only able to activate Flt-1 receptor. Based on the homology model, we also describe some specific interactions between VEGF and KDR.

## 2. Methodology

### 2.1. Construction of KDR receptor model

The primary sequence of the domain-2 of KDR ($KDR_{D2}$) of the human organism was extracted from *Swiss-Prot* protein data bank under the accession number P35968. The sequence identity between the domains-2 of Flt-1 and KDR is 40%. The structures of the copies of Flt-$1_{D2}$ complexed with VEGF (PDB code: 1FLT, 1.7 Å resolution) [13] were utilized as templates. Fig. 1 shows the sequence alignment of the second domains of Flt-1 and KDR obtained using ClustalW [16]. The sequence of $KDR_{D2}$ was folded onto the template sequence using the *Modeller* program [17], respecting the described alignment. Suitability of the generated structures was inspected using the *ProCheck* [18], *Profile-3D* [19] and *Prove* [20] computer programs.

```
                  10        20        30        40
          ....|....|....|....|....|....|....|....|
VEGFR1  1 GRPFVEMYSEIPEIIHMTEGRE--LVIPCRVTSPNITVTL 38
VEGFR2  1 RSPFIASVSDQHGVVYITENKNKTVVIPCLGSISNLNVSL 40

                  50        60        70        80
          ....|....|....|....|....|....|....|....|
VEGFR1 38 -KKFPLDTLIPDGKRIIWDSRKGFIISNATYKEIGLLTCE 77
VEGFR2 41 CARYPEKRFVPDGNRISWDSKKGFTIPSYMISYAGMVFCE 80

                  90
          ....|....|....|...
VEGFR1  78 ATVNGHLYK-TNYLTHRQ 94
VEGFR2  81 AKINDESYQSIMYIVVVV 98
```

**Fig. 1.** Sequence alignment of VEGF receptors Flt-1 (VEGFR-1) and KDR (VEGFR-2).

### 2.2. Construction of the complexes

In order to simplify the notation, from this point on the subscript $_{D2}$ on the receptor notation will be omitted, but we stress that only the domain-2 was taken into account in all analyses herein. The structures of the complexes VEGF/Flt-1 and PlGF/Flt-1 were previously determined and are available at the PDB under accession numbers 1FLT and 1RV6, respectively. Modeling of complexes VEGF/KDR, VEGFB/Flt-1, VEGFB/KDR and PlGF/KDR was achieved by structure superimposition. A similar process was performed before by Iyer and co-workers who generated the structure of the complex PlGF/Flt-1 before the determination of its crystal structure [11]. VEGF/KDR complex was generated by superimposing the KDR modeled structure onto the human VEGF/Flt-1 complex (PDB code: 1FLT) using the SwissPDB Viewer program [21]. VEGFB/KDR complex was generated by superimposing the KDR modeled structure and the structure of VEGFB (PDB code: 2C7W) onto the VEGF/Flt-1 complex. PlGF/KDR complex was generated by superimposing the KDR modeled structure onto the PlGF/Flt-1 complex (PDB code: 1RV6). VEGFB/Flt-1 complex was generated by superimposing the structure of VEGFB (PDB code: 2C7W) onto the VEGF/Flt-1 complex. Table 1 summarizes the construction of the models.

**Table 1**
Initial structures of the complexes.

| Complex | Origin of the factor | Origin of the receptor | Complex template | Residues of the factor/monomer | Residues of the receptor/copy |
|---|---|---|---|---|---|
| VEGF/Flt-1 | 1FLT[a] | 1FLT | – | 11-109 [99] | 132-226 [95] |
| VEGFB/Flt-1 | 2C7W[b] | 1FLT | 1FLT | 13-107 [95] | 132-226 [95] |
| PlGF/Flt-1 | 1RV6[c] | 1RV6 | – | 22-116 [95] | 132-223 [92] |
| VEGF/KDR | 1FLT | Model | 1FLT | 11-109 [99] | 122-220 [99] |
| VEGFB/KDR | 2C7W | Model | 1FLT | 13-107 [95] | 122-220 [99] |
| PlGF/KDR | 1RV6 | Model | 1RV6 | 22-116 [95] | 122-220 [99] |

The number of residues per chain is given inside the brackets.
[a] 1.7 Å resolution.
[b] 2.48 Å resolution.
[c] 2.45 Å resolution.

### 2.3. Initial structures of the unbound proteins

In addition to the complexes, molecular dynamics (MD) simulations were carried out also on the unbound systems. The analysis of unbound VEGF was performed using the MD trajectory obtained and described in our previous work [22]. The initial structures used for the simulations of VEGFB, PlGF e Flt-1 were taken from PDB files of codes 2C7W, 1RV6 and 1FLT, respectively. The initial structure used for the simulation of KDR was obtained by homology modeling (see above).

### 2.4. Molecular dynamics simulations

The unbound VEGF and the complex VEGF/Flt-1 were previously simulated and therefore we used the already generated trajectories for the analysis herein [22,23]. The other nine systems, unbound VEGFB, unbound PlGF, unbound Flt-1, unbound KDR, VEGF/KDR, VEGFB/KDR, PlGF/KDR, VEGFB/Flt-1 and PlGF/Flt-1, were simulated using the same protocol of our previous works [22,23].

All simulations were performed within the *NPT* ensemble, using the OPLS all atom force field [24] and the GROMACS [25,26] simulation package. Prior to the solvation procedure, energy minimizations using 1000 steps of steepest descent algorithm were performed in vacuum in order to remove high energy contacts. Systems were then solvated within rectangular boxes constructed with each edge from at least 9 Å of the protein surface. SPC water molecules [27,28] were used for this purpose. All systems were neutralized by replacing water molecules with $Na^+$ or $Cl^-$ counterions depending on the nature of the total charge of the proteins. The replacement of the water molecules with counterions was performed using the GROMACS program genion, which calculates the electrostatic potential before the insertion of each counterion in order to place it in the most favorable position. The factors VEGF, VEGFB and PlGF, which have total charges of $-12$, 0 and $-8$ (respectively), were neutralized with 12, 0 and 8 $Na^+$ counterions, respectively. The receptors Flt-1 and KDR, each of which having a total charge of +2, were neutralized with 2 $Cl^-$ counterions. Finally, the complexes VEGF/Flt-1, VEGF/KDR, VEGFB/Flt-1, VEGFB/KDR, PlGF/Flt-1 and PlGF/KDR which have total charges of -8, -8, 4, 4, -4, -4 (respectively), were neutralized with 8 $Na^+$, 8 $Na^+$, 4 $Cl^-$, 4 $Cl^-$, 4 $Na^+$, 4 $Na^+$ counterions, respectively. From this point on, periodic boundary conditions were used and long range electrostatic energies were computed using the PME method [29]. The solvated and neutralized systems were again submitted to energy minimization using 1000 steps of steepest descent algorithm, this time, within periodic boundary conditions. Simulations of solvent molecules and counterions were performed at 300 K and 1 atm for 40 ps in order to equilibrate them in the field of the potential exerted by the protein and to fill the cavities on the protein surface. In this procedure, the coordinates of the protein atoms were restrained with harmonic potentials. In all simulations, the bond stretching modes involving hydrogen atoms were constrained using the LINCS algorithm [30]. The equations of motion were integrated using the Verlet Leapfrog scheme [31]. Berendsen's formalism was used to control temperature and pressure with relaxation times of 0.1 and 1 ps, respectively. Isotropic coordinate scaling and isothermal compressibility of $4.5 \times 10^{-5}$ $bar^{-1}$ were used. [32] The molecular systems were again minimized for 1000 steps of steepest descent and then heated up to 300 K using 6 blocks of calculation in which the temperature was gradually increased. Each block lasted 50 ps, totalizing 300 ps of the heating process for each system. The equilibrated systems were simulated using a time-step of 0.002 ps. Except for the unbound receptors, for which simulations lasted for 20 ns, all other systems were simulated until achieving 10 ns.

## 3. Results and discussion

### 3.1. Structure and validation of modeled $KDR_{D2}$

The domain-2 of the KDR receptor ($KDR_{D2}$) can be classified as a member of the I-set of Ig-like domains. It consists of about 100 residues, which form two β-sheets that fold against each other to form a β-sandwich. Similarly to the domain-2 of Flt-1, one β-sheet is formed by strands A′, G, F, C, and the other is formed by strands B, E, and D. One disulfide bond is included in the hydrophobic core and connects strands B and F, stabilizing the fold of the domain [13,14].

The backbone conformation of the modeled structure was evaluated by using the PROCHECK program [18]. The analysis of the $\Psi/\Phi$ Ramachandran plot indicated that about 92% of the residues were located in the most favored regions (A, B, L). Six residues (6.9%) lied in generously allowed regions of the plot and only one residue (1.1%) had disallowed conformation (Asn205 which is the 84th residue of the model). Other stereochemical parameters for the KDR model were inside the expected values. Deviations of the atomic volumes from the standard values, evaluated as the volume *Z*-scores, are often used to assess the quality of protein crystal structures and were calculated using the PROVE [20] (PROtein Volume Evaluation) program. The average *Z*-score showed that the KDR model is well inside the range of highly resolved and well-refined protein crystal structures. The structure was further checked by Profile-3D [19] which gave a self-compatibility score of 32.57, higher than the template score of 26.66.

### 3.2. Structures of the modeled protein complexes

The structures of the complexes VEGF/KDR, VEGFB/Flt-1, VEGFB/KDR and PlGF/KDR were generated as previously described. No structure showed any atomic superposition and no high energy contact was observed after energy minimization. All complexes showed reasonable surface complementarity and consequently, a good fit was observed between the factors and the receptors.

Mutational studies revealed that some residues of VEGF which are important for KDR receptor binding are not those which are important for Flt-1 receptor binding [33]. This could be associated with a different orientation of VEGF in the complexes VEGF/Flt-1 and VEGF/KDR. However, the good surface complementarity observed in all complexes suggests that VEGF binds to both receptors with very similar orientation. This is quite intuitive as it would be hard to imagine two homologous proteins with a high degree of sequence identity binding to another protein with different orientations. Thus, reasons why some residues of VEGF are important to bind KDR and other residues are important to bind Flt-1 must be related to specific interactions, conformational adaptations, solvation and dynamical effects.

### 3.3. Stability of the trajectories

Molecular dynamics simulations were carried out as described in Section 2. Fig. 2 shows the time evolution of the RMSD of each protein system relative to initial structures. Trajectories of the unbound factors VEGF, VEGFB and PlGF were stable (Fig. 2a) over the 10 ns simulations and the average RMSD values are about ∼0.2 nm. Fig. 2b shows the RMSD of the receptors Flt-1 and KDR over the 20 ns simulation period. The trajectory of Flt-1 receptor was stable over the simulation with an average RMSD value of about ∼0.1 nm. In the case of KDR, whose starting point was the homology model, more pronounced conformational changes occurred until 13 ns of simulation and from this point on, the simulation was stable and converged to RMSD of about 0.3 nm. In

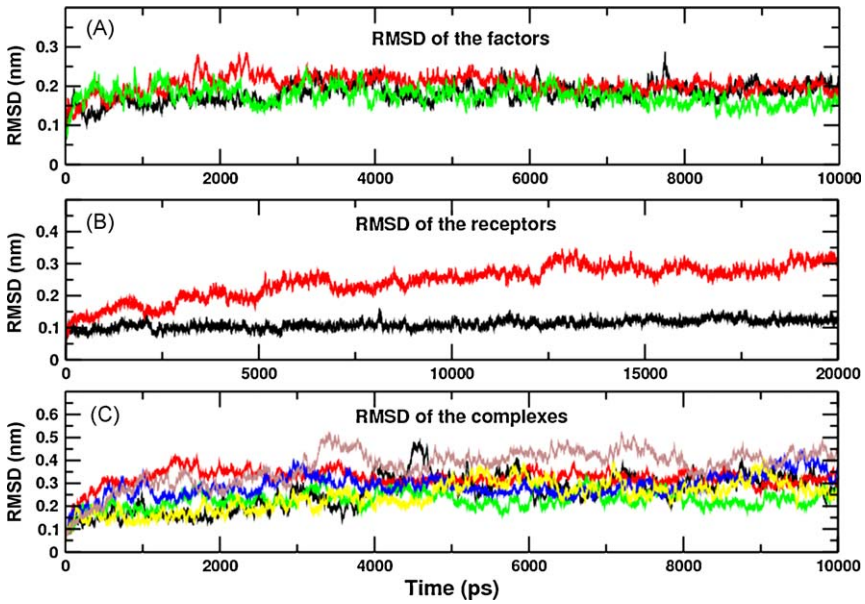


**Fig. 2.** RMSD time series of unbound factors, unbound receptors and complexes. (a) unbound factors: VEGF (black), VEGFB (red) and PlGF (green); (b) unbound receptors: Flt-1 (black) and KDR (red); (c) complexes: VEGF/Flt-1 (black), VEGF/KDR (red), VEGFB/Flt-1 (green), VEGFB/KDR (blue), PlGF/Flt-1 (yellow) and PlGF/KDR (brown).

the case of the complexes (Fig. 2c), all RMSD values are shown to be reasonably converged after 5 ns of simulation. The highest RMSD value (~0.4 nm) was shown by PlGF/KDR complex and the lowest value (~0.2 nm) was shown by VEGFB/Flt-1. The last 5 ns of each trajectory were considered as data acquisition time.

In order to show the structural changes in the complexes that occurred during the simulation in a more detailed way, RMSD values were averaged during the last 1 ns of simulation. The initial frame of each simulation was used as reference structure for the RMSD calculation. In Table 2, results are organized in five different types and each type is characterized by two groups. The first group consists of the set of atoms used for the least square fitting, and the second group consists of the set of atoms used in the RMSD calculation. For example, factor-Cα/factor-Cα means that the Cα-atoms of the growth factor were used for both fitting and RMSD calculation, whereas complex-Cα/complex-all means that all Cα-atoms of the complex were used for the fitting procedure, but the RMSD was calculated for all atoms of the complex. A look at Table 2 clearly shows that PlGF/KDR and VEGFB/KDR have higher RMSD values as a consequence of more pronounced conformational rearrangements. The large RMSD deviations found for these two complexes suggest that the factor-receptor interfaces contain unfavorable interactions, which tend to be minimized during the MD simulation (as a consequence of oversized repulsive forces) causing large conformational changes. All Flt-1 complexes, VEGF/Flt-1, VEGFB/Flt-1 and PlGF/Flt-1, are shown to not deviate significantly from their respective initial structures during the simulation. VEGF/KDR is the only KDR complex that does not show strong deviations from the initial structure, and surprisingly, it has RMSD values (considering the set that includes all atoms of the complexes and also considering the subset that includes the Cα atoms of the factors) even lower than those of VEGF/Flt-1 (e.g. 0.297 *vs.* 0.323 nm for the complexes and considering all Cα atoms; 0.188 *vs.* 0.215 nm for the factors and considering all Cα atoms), which initial coordinates were taken from a high resolution crystal structure.

### 3.4. Unbound KDR receptor: model vs. simulation structures

The model structure of the KDR receptor was utilized to build up all KDR complexes (VEGF/KDR, VEGFB/KDR and PlGF/KDR). However, for the performing of the subsequent analysis a molecular dynamics trajectory of the unbound KDR receptor in solution was necessary. As shown in Fig. 2b, the KDR receptor suffered conformational changes during the first 13 ns of simulation. Fig. 3 shows a superposition of the initial structure of the KDR receptor as obtained by homology modeling (shown in green) onto the last snapshot of the simulation (shown in brown). More pronounced conformational changes occur in the loop and terminal regions, which are the most flexible ones. The last snapshot of the simulation showed 81.6% of residues in the most favorable regions of the Ramachandran plot and 16.1% in additional favorable regions. No significant differences in the β-sheet arrangement were observed.

According to the preceding paragraph, an inspection of Fig. 2b shows that one of the unbound receptors (KDR) did not equilibrate

**Table 2**
RMSD averaged during the last 1 ns of simulation.

| Complex | Complex-Cα/ complex-Cα | Complex-Cα/ complex-all | Factor-Cα/ factor-Cα | Receptor-copy1-Cα/ receptor-copy1-Cα | Receptor-copy2-Cα/ receptor-copy2-Cα |
|---|---|---|---|---|---|
| VEGF/Flt-1 | 0.323 | 0.384 | 0.215 | 0.103 | 0.114 |
| VEGFB/Flt-1 | 0.219 | 0.313 | 0.179 | 0.117 | 0.120 |
| PlGF/Flt-1 | 0.311 | 0.385 | 0.158 | 0.105 | 0.154 |
| VEGF/KDR | 0.297 | 0.367 | 0.188 | 0.135 | 0.182 |
| VEGFB/KDR | 0.360 | 0.424 | 0.215 | 0.196 | 0.301 |
| PlGF/KDR | 0.409 | 0.492 | 0.166 | 0.194 | 0.213 |

The reference structure for the calculation of the RMSD was the first frame of the simulation. Results are organized in five different types and each type is characterized by two groups. The first group consists of the set of atoms used for the least square fitting, and the second group consists of the set of atoms used in the RMSD calculation.

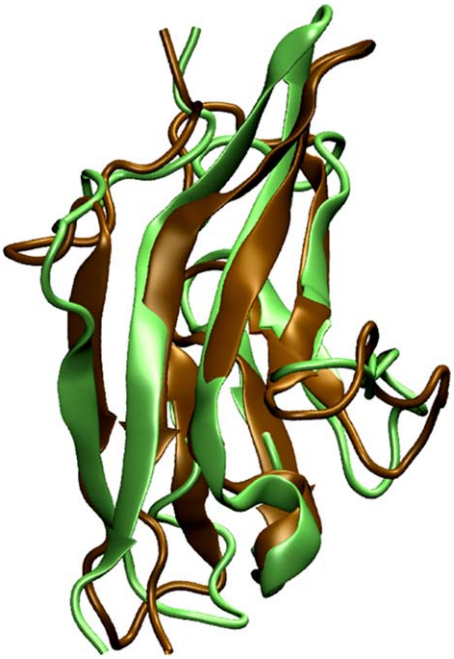

**Fig. 3.** Least square fitting superposition of KDR structures. In green: homology model (initial) structure; in brown: last snapshot of the simulation.

within the first 13 ns with reasonably large conformational changes taking place. This behavior is not totally surprising as the initial structure for the simulation corresponds to a theoretical prediction based on homology to a template. Possible causes for such rearrangements are: (i) molecular dynamics simulations depend strongly on the initial coordinates and velocities and even simulations with same initial coordinates and velocities can lead to different results if they are simulated in different machines with different precision due to numerical instabilities (of course, the magnitude of the differences are not expected to be large in this case); (ii) The modelled structure could be more similar to the template structure than it was supposed to be. In this case, the molecular dynamics simulation starts from a configuration that is not supposed to be close to the global minimum and starts sampling until a stable or meta-stable state is found, and a plateau region of the RMSD plot is consequently found. It is possible that this meta-stable state could be more similar to the yet unknown experimental structure. However, due to the limited sampling and inaccuracies of the force field this is not guaranteed and this structure could be even more different to the experimental structure than the initial structure of the simulation; (iii) For lucky, the homology model provided a structure that is very similar to the experimental structure, but the simulation diverged because of bad initial velocities or because this structure does not correspond to a minimum sufficiently deep in the potential energy surface (in terms of the force field Hamiltonian). (iv) The strong limitation of homology modelling on predicting the conformation of flexible parts (e.g. loops). In this case, a not so stable initial structure for the flexible part is subjected to MD simulations. The outcome of the simulations will probably lead to several conformations that in principle could actually exist as meta-stable states. As a consequence, the loop structure will be poorly defined, but the more rigid structures as beta-sheets and helices (if stable enough) could be somehow reliable.

Unfortunately, in this case, the only way of determining the best structure would be the experimental determination of the protein structure. However, the following sections illustrate a very good qualitative agreement in terms of the expected selectivity profile and also in terms of specific interactions between VEGF and KDR (in good agreement with alanine scanning mutagenesis).

### 3.5. Analysis of salt-bridge interactions

Fig. 4 shows the interfaces between the factor and the receptor at the last snapshot of the simulation of each complex. Charged residues are shown as a ball and stick representation. All complexes with Flt-1 show at least one salt-bridge interaction, with most of the charged residues that are close to the interface pointing toward the interface. The same occurs in the case of VEGF/KDR complex. On the other hand, it is not so obvious to visualize salt-bridges in the interfaces of the complexes VEGFB/KDR and PlGF/KDR. For these two complexes, the observed charged residues seem to be avoiding the interface. This occurs probably because the

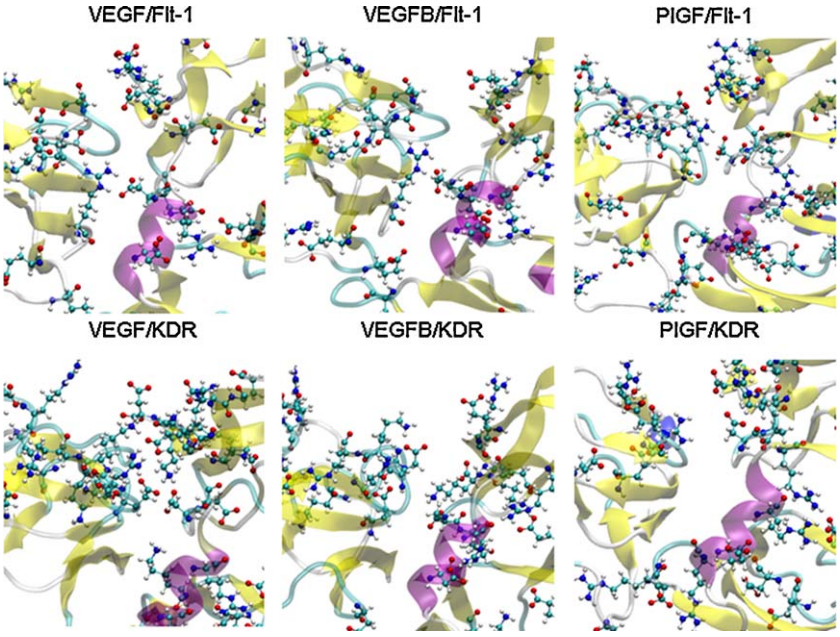


**Fig. 4.** Last snapshots of the simulations showing the interfaces factor/receptor. The backbone structures are shown as ribbon and colored by secondary structure. Charged residues are shown as ball and stick representation.

**Table 3**
Occupancies of salt-bridge interactions between factor and receptor.

| Salt-bridge | Occupancy |
|---|---|
| Complex: VEGF/Flt-1 | |
| Lys16/Glu141 | 0.753 |
| Asp63/Arg224 | 0.969 |
| Glu103/Lys171 | 0.387 |
| Complex: VEGF/KDR | |
| Lys16/Asp131 | 0.469 |
| Glu103/Arg164 | 0.728 |
| Glu103/Lys168 | 0.850 |
| Lys108/Glu140 | 0.276 |
| Complex: VEGFB/Flt-1 | |
| Asp63/Arg224 | 0.845 |
| Glu102/Lys200 | 0.270 |
| Complex: VEGFB/KDR | |
| Lys107/Glu140 | 0.137 |
| Glu102/Arg164 | 0.159 |
| Complex: PlGF/Flt-1 | |
| Arg96/Glu144 | 0.227 |
| Asp71/Arg224 | 0.500 |
| Glu111/Lys200 | 0.475 |
| Complex: PlGF/KDR | |
| Glu111/Lys168 | 0.911 |

charged residues in these complexes did not find a counterpart in the opposite side of the surface and prefer to be solvent exposed or to interact intramolecularly.

In order to show the differential salt-bridge formation in a more quantitative way we calculated the occurrence of all possible salt-bridges between factor and receptor. The existence of the salt-bridge was determined by a geometric criterion (distance cutoff of $<$0.5 nm between side chain heavy atoms). The last 5 ns of the trajectories were analyzed and results are shown in Table 3, which lists the salt-bridges with occupancy greater than 0.1.

In Table 3, one can observe that all Flt-1 complexes have a salt-bridge with Flt-1 residue Arg224, which is solvent exposed in the unbound Flt-1. The desolvation penalty originated by burying this charged residue in the complex interface must be somehow compensated. According to Table 3, the residue Asp63 of VEGF is the main responsible for mediating the interaction with Arg224. Indeed, experiments have shown that mutations of the negative residues Asp63, Glu64 and Glu67 of VEGF to alanine affect drastically the binding affinity for Flt-1 [33]. Probably, mutations to alanine of residues Asp63 of VEGFB and Asp71 of PlGF might cause similar effect on the binding affinity of VEGFB and PlGF for Flt-1. It is important to mention that the mutation of residues Asp63, Glu64 and Glu67 of VEGF do not cause a significant decrease in the binding affinity for KDR. This is also in agreement with our results, which show for the VEGF/KDR complex a very different pattern of salt-bridge interactions. In addition, a comparison of the KDR complexes shows an evident higher number of salt-bridges between KDR and VEGF (reflecting the expected trend being a consequence of the complementarity between polar interactions).

### 3.6. Hydrogen bonding analysis

Solute-solvent hydrogen bonds were used to estimate the role of desolvation upon the complex formation. The hydrogen bond was defined based on two geometric criteria: (i) a donor-acceptor distance cutoff (heavy atoms) of 0.35 nm and (ii) a donor-hydrogen-acceptor angle cutoff of 30°. This approach considers that the higher the number of hydrogen bonds between the protein and the solvent, the more solvated is the protein. Of course this is a very crude approach that depends on the system size, and number of hydrogen bond donors and acceptors on the protein surface. However, when considering proteins of the same size and considering only the difference between the unbound and bound states, it allows a comparative estimation of the desolvation penalty. Consider that $\Delta_{HB}$ is the variation of the number of hydrogen bonds with solvent molecules upon complex formation. In the case of the analyzed complexes, $\Delta_{HB}$ is given by the equation

$$\Delta_{HB} = HB_{complex\_solvent} - HB_{factor\_solvent} - 2HB_{receptor\_solvent} \quad (1)$$

where $HB_{complex\_solvent}$ is the average number of hydrogen bonds between the complex and the solvent, $HB_{factor\_solvent}$ is the average number of hydrogen bonds between the unbound factor and the solvent and $HB_{receptor\_solvent}$ is average the number of hydrogen bonds between the unbound receptor and the solvent. This equation varies with the composition of the protein complex. Herein, the term $HB_{receptor\_solvent}$ is multiplied by two because two molecules of the receptor are present in the complex. Table 4 shows the absolute average number of hydrogen bonds (averaged over the last 5 ns of each trajectory) between each system and the solvent and Table 5 shows the values of $\Delta_{HB}$ for each complex formation calculated by applying the values of Table 4 into Eq. (1).

**Table 4**
Absolute number of hydrogen bonds between each system with solvent molecules.

| System | Number of hydrogen bonds with solvent molecules |
|---|---|
| Unbound VEGF | 460 ± 24 |
| Unbound VEGFB | 426 ± 10 |
| Unbound PlGF | 462 ± 9 |
| Unbound Flt-1 | 228 ± 7 |
| Unbound KDR | 233 ± 8 |
| VEGF/Flt-1 | 915 ± 18 |
| VEGF/KDR | 918 ± 15 |
| VEGFB/Flt-1 | 861 ± 14 |
| VEGFB/KDR | 721 ± 40 |
| PlGF/Flt-1 | 843 ± 14 |
| PlGF/KDR | 736 ± 41 |

Both VEGF complexes, VEGF/Flt-1 and VEGF/KDR, show a small $\Delta_{HB}$ value in magnitude indicating that the number of hydrogen bonds with water molecules did not vary significantly upon complex formation. The small absolute values of $\Delta_{HB}$ suggest that the solvation penalty, in this case, is not a limiting factor for receptor binding process. VEGFB and PlGF complexes however show significant higher differences comparing to the complexes formed by Flt-1 and KDR receptors. VEGFB/Flt-1 and PlGF/Flt-1 have $\Delta_{HB}$ values of −21 and −75, respectively. In magnitude, these values are much lower than those calculated for VEGFB/KDR and PlGF/KDR, −171 and −192, respectively. This indicates that the solvation penalty is drastically higher for VEGFB/KDR and for the PlGF/KDR complexes.

Recently, Rodier et al. [34] have shown that most of the water molecules that are buried in the interface of protein–protein complexes are interacting with polar atoms via hydrogen bonds. The number of buried water molecules that directly contribute to

**Table 5**
Variation of the number of hydrogen bonds with solvent molecules upon complex formation.

| Complex | $\Delta_{HB}$ (variation of the number of H-bonds upon complex formation) |
|---|---|
| VEGF/Flt-1 | 915 − 460 − 456 = **−1** |
| VEGF/KDR | 918 − 460 − 466 = **−8** |
| VEGFB/Flt-1 | 861 − 426 − 456 = **−21** |
| VEGFB/KDR | 721 − 426 − 466 = **−171** |
| PlGF/Flt-1 | 843 − 462 − 456 = **−75** |
| PlGF/KDR | 736 − 462 − 466 = **−192** |

the interaction between the interfaces was shown to be on average about 10 and the number of water–protein polar interactions is on average about 28 per interface. Taking this data into account, one could try to figure out a rationale for the difference between the values of $\Delta_{HB}$ shown in Table 5. For example, one can compare the complexes VEGFB/Flt-1 and VEGFB/KDR, in which initial structures were generated by the same superposition method (i.e. to discard the possibility of the differences between the values of $\Delta_{HB}$ being artifacts generated by comparing structures obtained by superposition with those taken from crystal structure). If one considers that the complex VEGFB/Flt-1, which occurs in nature and is very tightly bound, has a $\Delta_{HB}$ of about −21, and that VEGFB/KDR, which does not occur in nature, has a $\Delta_{HB}$ of -171, one can hypothesize that: (i) VEGFB/KDR has a higher desolvation penalty compared to VEGFB/Flt-1. Therefore, favorable interactions are needed between VEGFB and KDR to overcome this penalty. (ii) The difference between the $\Delta_{HB}$ values might be related to the complementarity of polar interfaces. The mismatch of polar interactions between the proteins that form this complex can lead to a lower degree of "wetness" of the interface, i.e. the number of water molecules on the interface might be reduced because the polar atoms are being surrounded by hydrophobic ones. (iii) It is likely to imagine that nature could have selected minor structural and dynamical changes that expose to the solvent those polar atoms that are not in the interface to the bulk of the solvent upon complex formation. For example, the intramolecular salt bridge interaction between residues Arg82 and Glu42 of VEGF has a value of occupancy of 90.1% in the unbound protein and occupancy of 7.2% when VEGF is bound to KDR. A similar reduction occurs for the salt bridge between VEGF residues Lys74 and Glu34 (94.5% occupancy in the unbound VEGF against 40.5% occupancy in VEGF/KDR). The disruption of these salt bridges allows the charged residues to interact with the bulk of the solvent, increasing the number of solute-solvent hydrogen bonds upon complex formation and reducing the solvation penalty.

Table 6 shows the average number of hydrogen bonds between the factor and the two copies of the receptor. The complexes VEGFB/KDR and PlGF/KDR show the smallest values suggesting that donor/acceptor hydrogen bond groups in these protein interfaces are less complementary as compared to the protein complexes VEGF/Flt-1, VEGF/KDR, VEGFB/Flt-1 and PlGF/Flt-1.

### 3.7. Interface area

The overall SASA was computed for each system. Calculations of SASA were performed using the *NACCESS* computational code with default van der Waals atomic radii, a probe radius of 1.4 Å, and a slice thickness of 0.05 Å. The interface area between the factor and the two copies of the receptor of each complex was calculated using Eq. (2) [35].

$$S_{int} = \frac{1}{2}SASA_{factor} + 2SASA_{receptor} - SASA_{complex} \quad (2)$$

where $S_{int}$ is the interface area and SASA is the solvent accessible surface area. The subscripts define the respective system.

**Table 6**
Number of hydrogen bonds between factor and receptor.

| Complex | Number of H-bonds between factor and receptor |
|---|---|
| VEGF/Flt-1 | 10 ± 1 |
| VEGF/KDR | 10 ± 2 |
| VEGFB/Flt-1 | 8 ± 1 |
| VEGFB/KDR | 5 ± 2 |
| PlGF/Flt-1 | 8 ± 2 |
| PlGF/KDR | 6 ± 1 |

**Table 7**
Percentage of buried area of each receptor by each factor.

| | %total BA | Polar/Apolar | %Polar | %Apolar |
|---|---|---|---|---|
| BA of Flt-1 by VEGF | 28.5 | 0.44 | 8.7 | 19.7 |
| BA of KDR by VEGF | 36.9 | 0.40 | 10.5 | 26.4 |
| BA of Flt-1 by VEGFB | 25.2 | 0.49 | 8.2 | 16.9 |
| BA of KDR by VEGFB | 24.0 | 0.36 | 6.3 | 17.6 |
| BA of Flt-1 by PlGF | 31.4 | 0.48 | 10.1 | 21.2 |
| BA of KDR by PlGF | 32.9 | 0.36 | 8.6 | 24.2 |

**Table 8**
Percentage of buried area of each factor by each receptor.

| | %total BA | Polar/Apolar | %Polar | %Apolar |
|---|---|---|---|---|
| BA of VEGF by Flt-1 | 28.0 | 0.44 | 8.6 | 19.4 |
| BA of VEGF by KDR | 41.9 | 0.40 | 12.0 | 30.0 |
| BA of VEGFB by Flt-1 | 27.7 | 0.49 | 9.1 | 18.6 |
| BA of VEGFB by KDR | 30.5 | 0.36 | 8.0 | 22.4 |
| BA of PlGF by Flt-1 | 34.8 | 0.48 | 11.3 | 23.5 |
| BA of PlGF by KDR | 42.1 | 0.36 | 11.1 | 31.0 |

The values of SASA were obtained by averaging over the last 5 ns of each trajectory. Once the interface areas were calculated, it was possible to calculate the buried areas by subtracting the interface area from the total surface areas of the individual molecules. The percentage of total, polar and apolar buried areas (BA) are shown below for the factors (Table 7) and for the receptors (Table 8). The sum %Polar + %Apolar is equal to %total BA. For example, 28.5% (8.7% polar + 19.7% apolar) of the total area of Flt-1 is buried by VEGF. The Polar/Apolar ratio provides evaluation of the polarity of the interface. It is interesting to note that the complexes which are not formed (VEGFB/KDR and PlGF/KDR) have a significantly lower Polar/Apolar ratio. This could be attributed to a mismatch of polar interactions in the non-favored complexes.

### 3.8. Interactions between the domain-2 of KDR and VEGF

The model of VEGF/KDR complex presented herein has made possible a structural description of the interactions formed between VEGF and KDR. Experimental data from alanine scanning mutagenesis show that VEGF residues Phe-17, Tyr-21, Ile-43, Ile-46 Glu-64, Gln-79 and Ile-83 are the most important for KDR binding [6,33]. In this section we describe how these residues interact with KDR (Fig. 5).

Analyses of the MD trajectories have shown that the VEGF residues Phe-17 and Tyr-21 interact with the KDR residues Tyr-165, Met-197, Ile-215 and Met-213. These interactions are mainly hydrophobic and no direct hydrogen bonds are observed between these residues (Fig. 5a). The residue Ile-43 (of VEGF loop-1) does not show any short range interaction with the domain-2 of KDR, and probably interacts only with domain-3. On the other hand, Ile-46 is shown to interact with domain-2 residues Val-217 and Val-219 (Fig. 5b). The residue Glu-64 (of VEGF loop-2) and the residue Ile-43, do not show any short range interaction to domain-2. It is important to mention that experimental deletion studies of the Flt-1 and KDR receptors have shown that domains 2-3 are sufficient for binding VEGF with near wild-type affinity [13], and that the additional deletion of domain-3 of Flt-1 only causes a 20-fold decrease in binding affinity as compared to a larger than 1000-fold decrease in the case of KDR [13]. This implies that the relative importance of domain-3 is different between these receptors. Alanine scanning of the receptor-binding site of VEGF was used to measure the relative importance of individual VEGF residues to Flt-1 and KDR constructs containing domains 1-3 [33]. Almost all residues of VEGF that are important for Flt-1 binding are located in

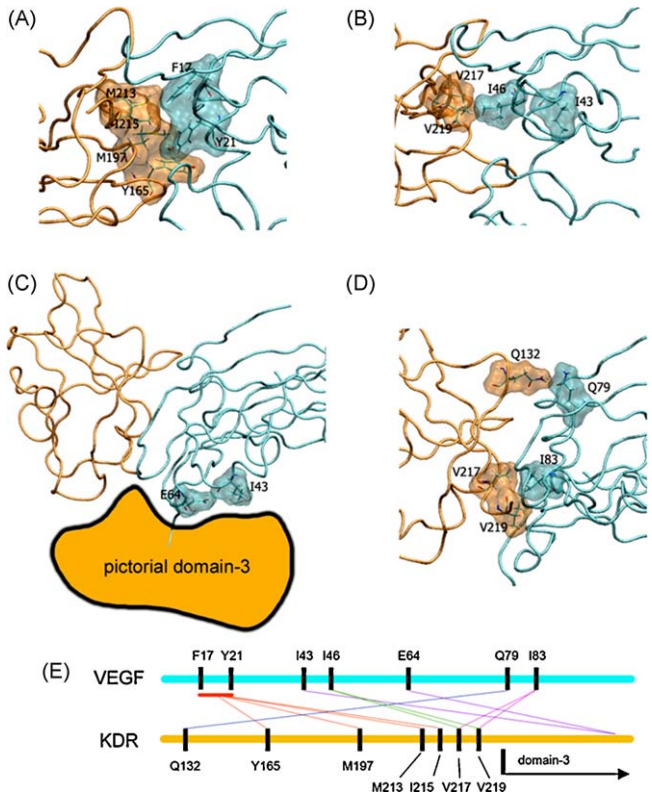


**Fig. 5.** Interactions between VEGF and KDR receptor. VEGF is shown in cyan ribbon and KDR in orange ribbon. Side-chains of highlighted amino acids are depicted as sticks and contoured with transparent surfaces. (a) N-terminal residues of VEGF Phe17 and Tyr21 are shown to interact with Tyr165, Met197, Met213 and Ile215 of KDR. (b) Loop-1 residues of VEGF Ile43 and Ile46. Only Ile46 interacts with domain-2 of KDR. (c) Pictorial representation of domain-3 of KDR and its proximity to residues Glu64 and Ile43. (d) Loop-3 residue Gln79 interacts with Gln132 of KDR and loop-3 residue Ile83 interacts with Val217 and Val219 of KDR. (e) Simplified representation of the interactions between VEGF and KDR.

the interface with the domain-2. The same does not occur for the residues that are important for KDR binding as some of these residues are located outside the contact region with domain-2. Wiesmann et al. [13] suggested based on the older alanine scanning studies of Muller et al. [6] that the VEGF residues Glu-64 and Ile-43 would possibly interact with the domain-3. In agreement with these ideas, we propose a pictorial model for the orientation of the domain-3 of KDR (Fig. 5c). The model illustrates a situation in which the residues Glu-64 and Ile-43 are directly and exclusively interacting with this domain. Fig. 5d shows the interactions of the residues Gln-79 and Ile-83 (of VEGF loop-3) with KDR. They are shown to be far apart to each other and interacting with different parts of domain-2. Gln-79 interacts via hydrogen bond with Gln-132 of KDR. Ile-83 interacts with C-terminal Val-217 and Val-219 of KDR. Finally, Fig. 5e shows a schematic representation (in terms of primary sequence) of the described interactions.

### 3.9. Convergence of eigenvectors from principal component analysis

For all different systems considered in this work (unbound VEGF, VEGFB and PlGF; unbound receptors Flt-1 and KDR; and complexes VEGF/Flt-1, VEGF/KDR, VEGFB/Flt-1, VEGFB/KDR, PlGF/Flt-1 and PlGF/KDR) the covariance matrices were constructed and diagonalyzed several times in order to obtain the convergence of the root mean square inner products (RMSIP) within the subspace consisting of the first 10 eigenvectors. A similar analysis was performed before for unbound VEGF [22] and for VEGF bound to Flt-1 and VEGF bound to a peptide inhibitor [23].

In general, the RMSIP analysis is used to validate a principal component analysis PCA analysis. For example, it was used to validate the essential dynamics of protein L and Cytochrome c551 [36]. Herein, we calculate RMSIP as a probe for the convergence of the modes of essential dynamics (essential space being formed by the first 10 eigenvectors). For the calculation of the RMSIP we consider the overlap between eigenvectors collected from subparts of increasing time lengths calculated from two halves of the last 5 ns of each trajectory. In order to avoid correlations, we only calculate overlaps between subparts of the first half with subparts of the second half of the trajectories. Solvent and counterions were excluded from the analysis and the overall translational and rotational motions were suppressed by prior least square fitting superposition. It is also important to mention that only Cα atoms were considered in this analysis and that the first 5 and last 5 residues of each chain were excluded in order to minimize the artifacts introduced by the highly flexible motion of the terminal residues. In Fig. 6 we show the overlaps obtained from pairs of subparts ranging from 5 to 2500 ps (full half).

In Fig. 6a, we show the analysis for the VEGF residues. The black line corresponds to the VEGF in its unbound state, while the red and green line corresponds to VEGF from the simulation in which it is complexed with Flt-1 and KDR, respectively. Basically, we are interested in the effect of binding to one of these receptors in the convergence of the dynamical modes of VEGF. The modes are reasonably converged for the 3 systems. The unbound VEGF and the VEGF in the presence of KDR are converged to a RMSIP value of about 0.5 and the VEGF in the presence of Flt-1 is converged to a slightly smaller RMSIP value of 0.44.

In Fig. 6b, we show the results regarding the VEGFB residues. The black line corresponds to the unbound VEGFB, the red line to VEGFB in the presence of Flt-1 and the green line to VEGFB in the presence of KDR. In this case, the unbound and the one in the presence of Flt-1 are better converged to a RMSIP value about 0.45. The VEGFB in the presence of KDR converges fast (∼250 ps) but to a maximum value of 0.32.

In Fig. 6c, we show a similar analysis for PlGF. In this case, the unbound PlGF converges to a higher value than both bound cases. However, the difference here is rather small, unbound PlGF with RMSIP of about 0.46 and the bound states about 0.42.

The same analysis was performed for the receptors in the unbound form or in the presence of the growth factors. In Fig. 6d, we show that the unbound Flt-1 (black line) converges to a RMSIP of about 4.2, followed by Flt-1 in the presence of VEGF (red line) with RMSIP about 3.9, then Flt-1 in the presence of VEGFB (green line) with RMSIP about 3.7 and Flt-1 in the presence of PlGF (blue line) with RMSIP about 3.5.

In Fig. 6e, the same is showed for receptor KDR. It is interesting to note that the unbound KDR and the KDR in the presence of VEGF converge to almost the same RMSIP value of 3.9, while KDR in the presence of VEGFB and KDR in the presence of PlGF show RMSIP values of about 0.29 and 0.26, respectively. This probably indicates that the dynamical modes of KDR are not reasonably converged in the simulations in which KDR is in the complex with VEGFB and PlGF.

Finally, Fig. 6f, g, and h shows the results for the entire complexes. In Fig. 6f, the complexes involving VEGF (VEGF/Flt-1 and VEGF/KDR) are shown. Both VEGF complexes converge to the same RMSIP value of about 4.8. In Fig. 6g, the results for the complexes involving VEGFB are shown. VEGFB/Flt-1 (black line) is converged to a RMSIP value over 5.6, while VEGFB/KDR is converged to RMSIP of about 3.9. In Fig. 6h, the results for the complexes involving PlGF are shown. PlGF/Flt-1 is converged to RMSIP of about 4.7 and PlGF/KDR to about 3.5.

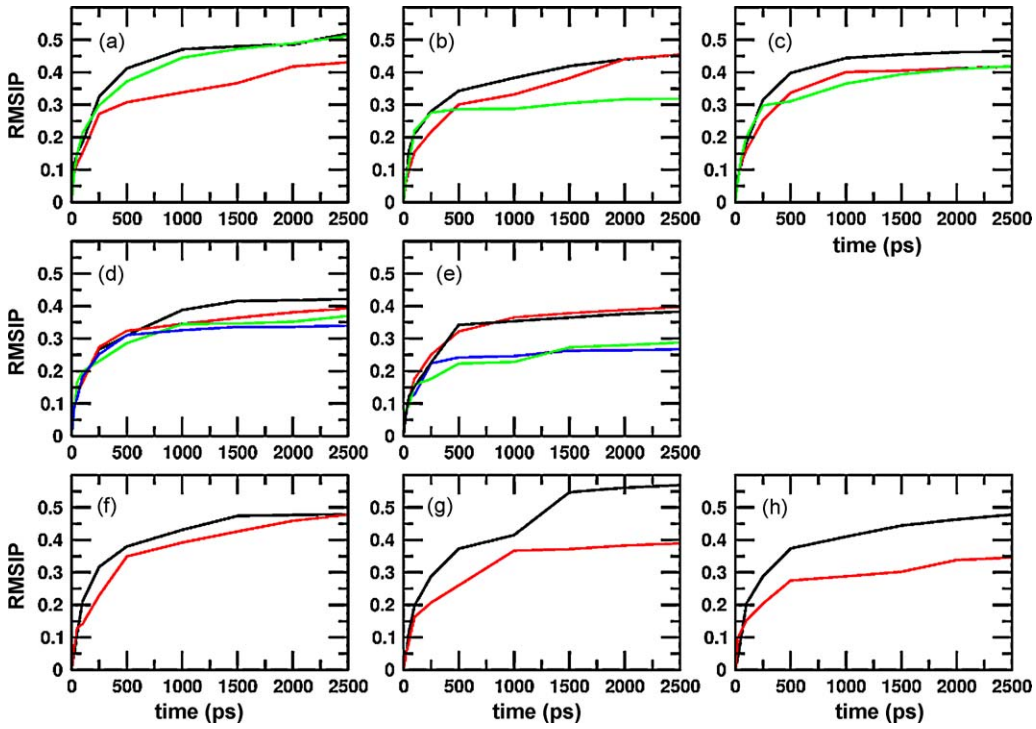


**Fig. 6.** Root mean square inner products (RMSIP) of eigenvectors obtained from two different time windows (of 2500 ps), considering the subspace which includes the first 10 eigenvectors obtained from principal component analysis. (a) Unbound VEGF (black); VEGF in the complex with Flt-1 (red); and VEGF in the complex with KDR (green). (b) Unbound VEGFB (black); VEGFB in the complex with Flt-1 (red); and VEGFB in the complex with KDR (green). (c) Unbound PlGF (black); PlGF in the complex with Flt-1 (red); PlGF in the complex with KDR (green). (d) Unbound Flt-1 (black); Flt-1 in the complex with VEGF (red); Flt-1 in the complex with VEGFB (green); Flt-1 in the complex with PlGF (blue). (e) Unbound KDR (black); KDR in the complex with VEGF (red); KDR in the complex with VEGFB (green); KDR in the complex with PlGF (blue). (f) Complex VEGF/Flt-1 (black); complex VEGF/KDR (red). (g) Complex VEGFB/Flt-1 (black); complex VEGFB/KDR (red). (h) Complex PlGF/Flt-1 (black); complex PlGF/KDR (red).

The values of RMSIP for the complexes VEGFB/KDR and PlGF/KDR are systematically lower compared to those of VEGFB/Flt-1 and PlGF/Flt-1. Possibly, this indicates that the dynamical modes of the considered essential subspace are not reasonably converged for the VEGFB/KDR and PlGF/KDR complexes. The poor convergence in this case cannot simply be related to the quality of the KDR model because for the complex VEGF/KDR a much higher RMSIP value is found. Therefore, the bad convergence of the RMSIP values observed for the complexes VEGFB/KDR and PlGF/KDR indicates that dynamical and conformational changes are still taking place in the level of backbone.

### 3.10. Overall structure and dynamics of the protein complexes

In the present work, we propose structural models and perform molecular dynamics simulations in order to learn how receptor selectivity is achieved by growth factors. At this point, we discuss the reliability of the presented structural models and the time length of the simulations.

In the present study we try to address the selective activation of vascular growth factor receptors. The complex factor/receptor recognition and activation is modulated by subtle differences between the structures of the proteins. For many of these proteins, no experimental structure is available and therefore, the use of modeled structures and computer simulation is appealing. If the structures of the modeled complexes are reliable, the analysis on the level of the amino acid interactions could be able to provide insight on the selectivity profile. In 2001, Iyer et al. [11] used the superposition method in order to propose a model for the interaction of PlGF and Flt-1 receptor. Three years later, the crystal structure of the complex was determined and showed that the binding mode proposed before was in a very good agreement. Considering that the known crystal structures of vascular growth factors show only small differences between protein backbones, and that the alanine scanning experiments are also consistent to very similar binding mode, it is to be expected that the proposed models for the unbound KDR receptor and for the protein complexes represent the correct binding orientation and capture most of the important interactions. However, amino acid side chains are known to be very flexible when they are solvent exposed, and therefore, it might be that some disagreements with yet undetermined crystal structures could take place. Nevertheless, the qualitative results shown herein are consistent with the selectivity profile, which may indicate that the protein–protein interface of the obtained model complexes show a reasonable adaptation of the side chains. In addition, we want to stress that, for the complexes VEGFB/KDR and PlGF/KDR, which are not stable at physiological conditions, more rearrangements could have occurred in order to minimize the energy in the protein–protein interface, caused probably by strong mismatch of polar interactions in these complexes. One could investigate these rearrangements by looking at Fig. 2 and Table 2, which show larger RMSD values for the backbone of the complexes VEGFB/KDR and PlGF/KDR. The convergence of the eigenvectors of principal component analysis is also in perfect agreement with these observations (Fig. 6).

The molecular dynamics simulations showed initial structure rearrangements for the complexes, which occurred within the first 5 ns. The most pronounced changes occurred in the loops and terminal regions. Most of the loop residues of VEGF are known to be at the complex interface, and probably, some of the loop rearrangements were caused by the accommodation of the interface residues to the potential exerted by the receptor. The simulation time of few ns was shown to be enough for a coarse equilibration of the interface. The RMSD plots reach a plateau after 5 ns (Fig. 2), which may indicate a meta-stable state. We stress that

the fact that RMSD reached a plateau does not guarantee that other properties of the system, such as hydrogen bonds, SASA and salt-bridges are necessarily (fully) converged. In addition, motions of higher amplitude, partial unfolding such as the loss of secondary structure content, and complex dissociation are expected to occur in a much longer time-scale for proteins of this size. Therefore, our results and conclusions are limited by our observations in the time-scale of few nanoseconds. One possible outcome of the extension of the trajectories to hundreds of nanoseconds is the partial dissociation of the complexes VEGFB/KDR or PlGF/KDR. As we showed, a big desolvation penalty and unfavorable interactions are observed for these complexes. The higher values of RMSD (Table 2) could be also an indicative of future dissociation, change of orientation or partial loss of secondary structure.

## 4. Conclusions

Many biological processes are governed by specific or highly selective protein–protein recognition. The case of vascular endothelial growth factors exemplifies how nature can select different biological responses by expressing homologous proteins, which in shape are very similar but show subtle differences that are able to regulate selectivity or promiscuity regarding different receptors.

In this work, we showed that the polarity of the interface, the presence or absence of salt-bridge interactions, and the variation of the number of hydrogen bonds upon complex formation are consistent with the experimental observed selectivity profile (VEGF activates receptors Flt-1 and KDR but PlGF and VEGFB only activate receptor Flt-1). Considering altogether, the data showed in this work indicate that selectivity seems to be a consequence of the complementarity of polar interactions. We show that, in comparison to the other examined complexes presented herein, VEGFB/KDR and PlGF/KDR are outliers as the properties calculated for these complexes are numerically different from the others. The VEGFB/KDR and PlGF/KDR complexes are not experimentally observed probably due to a strong repulsion caused by a mismatch of the polar protein–protein interactions which make impossible to compensate for a large dessolvation penalty.

## Acknowledgements

The authors would like to thank Dr. Philippe Hünenberger for fruitful suggestions on the hydrogen bond analysis and Dr. Jožica Dolenc for comments on the manuscript. Brazilian National agencies CNPq and FAPERJ are also acknowledged for financial support.

## References

[1] N. Ferrara, H.P. Gerber, J. LeCouter, The biology of VEGF and its receptors, Nature Medicine 9 (2003) 669–676.
[2] N. Ferrara, T. Davis-Smyth, The biology of vascular endothelial growth factor, Endocrine Review 18 (1997) 4–25.
[3] J. Folkman, Angiogenesis, Annual Review of Medicine 57 (2006) 1–18.
[4] P. Carmeliet, Angiogenesis in life, disease and medicine, Nature 438 (2005) 932–936.
[5] X.R. Li, U. Eriksson, Novel VEGF family members: VEGF-B, VEGF-C and VEGF-D, International Journal of Biochemistry & Cell Biology 33 (2001) 421–426.
[6] Y.A. Muller, B. Li, H.W. Christinger, J.A. Wells, B.C. Cunningham, A.M. DeVos, Vascular endothelial growth factor: Crystal structure and functional mapping of the kinase domain receptor binding site, Proceedings of the National Academy of Sciences of the United States of America 94 (1997) 7192–7197.
[7] A.K. Olsson, A. Dimberg, J. Kreuger, L. Claesson-Welsh, VEGF receptor signalling—in control of vascular function, Nature Reviews Molecular Cell Biology 7 (2006) 359–371.
[8] M.G. Achen, B.K. McColl, S.A. Stacker, Focus on lymphangiogenesis in tumor metastasis, Cancer Cell 7 (2005) 121–127.
[9] Y. Yamazaki, T. Morita, Molecular and functional diversity of vascular endothelial growth factors, Molecular Diversity 10 (2006) 515–527.
[10] Y.A. Muller, H.W. Christinger, B.A. Keyt, A.M. deVos, The crystal structure of vascular endothelial growth factor (VEGF) refined to 1.93 angstrom resolution: multiple copy flexibility and receptor binding, Structure 5 (1997) 1325–1338.
[11] S. Iyer, D.D. Leonidas, G.J. Swaminathan, D. Maglione, M. Battisti, M. Tucci, M.G. Persico, K.R. Acharya, The crystal structure of human placental growth factor-1 (PlGF-1) an angiogenic protein, at 2.0 Å resolution, Journal of Biological Chemistry 276 (2001) 12153–12161.
[12] S. Iyer, P.D. Scotney, A.D. Nash, K.R. Acharya, Crystal structure of human vascular endothelial growth factor-B: Identification of amino acids important for receptor binding, Journal of Molecular Biology 359 (2006) 76–85.
[13] C. Wiesmann, G. Fuh, H.W. Christinger, C. Eigenbrot, J.A. Wells, A.M. deVos, Crystal structure at 1,7 angstrom resolution of VEGF in complex with domain 2 of the Flt-1 receptor, Cell 91 (1997) 695–704.
[14] M.A. Starovasnik, H.W. Christinger, C. Wiesmann, M.A. Champe, A.M. de Vos, N.J. Skelton, Solution structure of the VEGF-binding domain of Flt-1: Comparison of its free and bound states, Journal of Molecular Biology 293 (1999) 531–544.
[15] H.W. Christinger, G. Fuh, A.M. de Vos, C. Wiesmann, The crystal structure of placental growth factor in complex with domain 2 of vascular endothelial growth factor receptor-1, Journal of Biological Chemistry 279 (2004) 10382–10388.
[16] D. Higgins, J. Thompson, T. Gibson, J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Research 22 (1994) 4673–4680.
[17] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, Journal of Molecular Biology 234 (1993) 779–815.
[18] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, Journal of Applied Crystallography 26 (1993) 283–291.
[19] R. Luthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, Nature 356 (1992) 83–85.
[20] J. Pontius, J. Richelle, S.J. Wodak, Deviations from standard atomic volumes as a quality measure for protein crystal structures, Journal of Molecular Biology 264 (1996) 121–136.
[21] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, Electrophoresis 18 (1997) 2714–2723.
[22] B.A.C. Horta, J.J.V. Cirino, R.B. de Alencastro, Dynamical behavior of the vascular endothelial growth factor: biological implications, Proteins 67 (2007) 517–525.
[23] B.A.C. Horta, J.J.V. Cirino, R.B. De Alencastro, On the structure, interactions, and dynamics of bound VEGF, Journal of Molecular Graphics and Modelling 26 (2008) 1091–1103.
[24] G.A. Kaminski, R.A. Friesner, J. Tirado-Rives, W.L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides, Journal of Physical Chemistry B 105 (2001) 6474–6487.
[25] H.J.C. Berendsen, D. Vanderspoel, R. Vandrunen, Gromacs—a message-passing parallel molecular-dynamics implementation, Computer Physics Communications 91 (1995) 43–56.
[26] E. Lindahl, B. Hess, D. van der Spoel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, Journal of Molecular Modeling 7 (2001) 306–317.
[27] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, J. Hermans, Interaction models for water in relation to protein hydration, in: B.E. Pullman (Ed.), Intermolecular Forces, Dordrecht, the Netherlands, D. Reidel Publishing Company, Dordrecht, 1981, pp. 331–342.
[28] C.D. Berweger, W.F. Vangunsteren, F. Mullerplathe, Force-field parametrization by weak-coupling—reengineering SPC water, Chemical Physics Letters 232 (1995) 429–436.
[29] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, Journal of Chemical Physics 103 (1995) 8577–8593.
[30] B. Hess, H. Bekker, H.J.C. Berendsen, J. Fraaije, LINCS: A linear constraint solver for molecular simulations, Journal of Computational Chemistry 18 (1997) 1463–1472.
[31] R.W. Hockney, The potential calculation and some applications, Methods in Computational Physics 9 (1970) 136–211.
[32] H.J.C. Berendsen, J.P.M. Postma, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, Journal of Chemical Physics 81 (1984) 3684–3690.
[33] B. Li, G. Fuh, G. Meng, X. Xin, M.E. Gerritsen, B. Cunningham, A.M. De Vos, Receptor-selective variants of human vascular endothelial growth factor, Journal of Biological Chemistry 275 (2000) 29823–29828.
[34] F. Rodier, R.P. Bahadur, P. Chakrabarti, J. Janin, Hydration of protein–protein interfaces, Proteins 60 (2005) 36–45.
[35] P.J. Kundrotas, E. Alexov, Electrostatic properties of protein–protein complexes, Biophysical Journal 91 (2006) 1724–1736.
[36] A. Amadei, M.A. Ceruso, A. Di Nola, On the convergence of the conformational coodinates basis set obtained by the essential dynamics analysis of proteins' molecular dynamics simulations, Proteins 36 (1999) 419–424.