Journal of Molecular Graphics and Modelling 45 (2013) 122–127

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

# The 'order-to-disorder' conformational transition in CD44 protein: An umbrella sampling analysis

Wojciech Plazinski [a,*], Agnieszka Knys-Dzieciuch [b]

[a] J. Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, ul. Niezapominajek 8, 30-238 Krakow, Poland
[b] Department of Theoretical Chemistry, Faculty of Chemistry, M. Curie-Sklodowska University, pl. M. Curie-Sklodowskiej 3, 20-031 Lublin, Poland

ARTICLE INFO

*Article history:*
Received 12 February 2013
Received in revised form 16 July 2013
Accepted 6 August 2013
Available online 19 August 2013

*Keywords:*
CD44
Steered molecular dynamics
Conformational transition
Free energy profiles
Umbrella sampling
Glycoproteins

ABSTRACT

The molecule of CD44, a membrane protein being the major cell surface receptor for hyaluronan, is postulated to undergo the conformational rearrangement called the 'order-to-disorder' transition. The experimental studies suggest that the Tyr161 residue is crucial for maintaining the equilibrium between the 'ordered' (O) and 'partially disordered' (PD) forms of CD44. The molecular modeling study based on the umbrella sampling protocol was carried out separately for the wild-type CD44 and Tyr161Ala mutant in order to gain more insight into the molecular mechanism of the O-PD transition and to clarify the role of the Tyr161 amino acid residue. The calculated free energy profiles associated with the initial stages of the O-PD conformational transition allow to identify the crucial steps of this process and their molecular details. The differences between the wild-type CD44 and the Tyr161Ala mutant are very insignificant which allows for speculating that, surprisingly, the role of Tyr161 in the O-PD transition is not connected with disrupting the attractive Glu48-Tyr161 and Leu52-Tyr161 interactions but with other types of (unknown yet) interactions located outside the β7-β8 loop or with the existence of the PD-like structure in which the terminal lobe remains located under the β7-β8 loop.



## 1. Introduction

CD44 is a principal cell surface receptor for hyaluronan (HA) and plays an important role in a wide variety of biological and pathological events. The experimental studies let to postulate that the CD44 molecule can undergo the conformational changes which switch the receptor 'on' or 'off' under appropriate circumstances [1]. According to Ref. [2], such conformational type switching may be induced by the binding of HA but the subsequent studies [1] showed that large rearrangements in the CD44 molecule can occur either in the presence or the absence of ligand. Our previous study reported in Ref. [3] was based on the structures of CD44 in the complex with the HA heptamer and resolved by Banerji et al. [4]. Two conformational forms (called 'A' and 'B') of the CD44/HA complex were identified, differing in orientation of a crucial HA-binding residue (Arg45). Thus, it was suggested that there is an interconversion between these two conformers upon ligand (i.e. HA) binding. One can also expect that the 'B' form represents a higher-affinity ligand-bound state as it corresponds to a more intimate contact between the HA chain and CD44. One of the results of our previous investigations [3] was the conclusion that the structures corresponding to the two crystal forms ('A' and 'B', according to Ref. [4]) can be distinguished only in some cases when considering the values of dihedral angle $\varphi$ at the Tyr46 residue (this free energy barrier is, however, not high and can be overcome several times during the unbiased molecular dynamics simulation lasting ~100 ns). On the other hand, the free energy barrier associated with the Arg56-HA distance (i.e. the main parameter varying between 'A' and 'B') is much lower and, as a result, both 'A' and 'B' states reduce to an average, dynamic structure. Such behavior was observed for both the liganded and unliganded forms of CD44.

Apart from summarizing our previous results, we would like to comment on the very closely related studies reported by Jamison et al. The paper [5] was unknown to us during publishing our results collected in [3] (thus, it was not mentioned there) but the obtained results deserve more profound comparison with our data. The authors concluded [5] that there exists significant (~6.5 kcal/mol) free energy barrier associated with the changing Arg56-HA distance; furthermore, this barrier is correlated with the varying value of the dihedral angle $\varphi$ at the Tyr46 residue which was identified as molecular switch governing the HA-CD44 interaction strength. In the case of our study, the $\varphi$-related free energy barrier was not correlated with the Arg45-HA distance (see Fig. 2 in Ref. [3]). Due to the similarities in both studies (e.g. using the same types of structures, computational methods, etc.), such contradictory results are most likely caused by using different potentials (force fields) for the interactions in the system. Currently, it is hard to judge which

* Corresponding author. Tel.: +48 815375685; fax: +48 815375685.
*E-mail addresses:* wojtek_plazinski@o2.pl, wojtek@vega.umcs.lublin.pl (W. Plazinski).

1093-3263/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2013.08.002

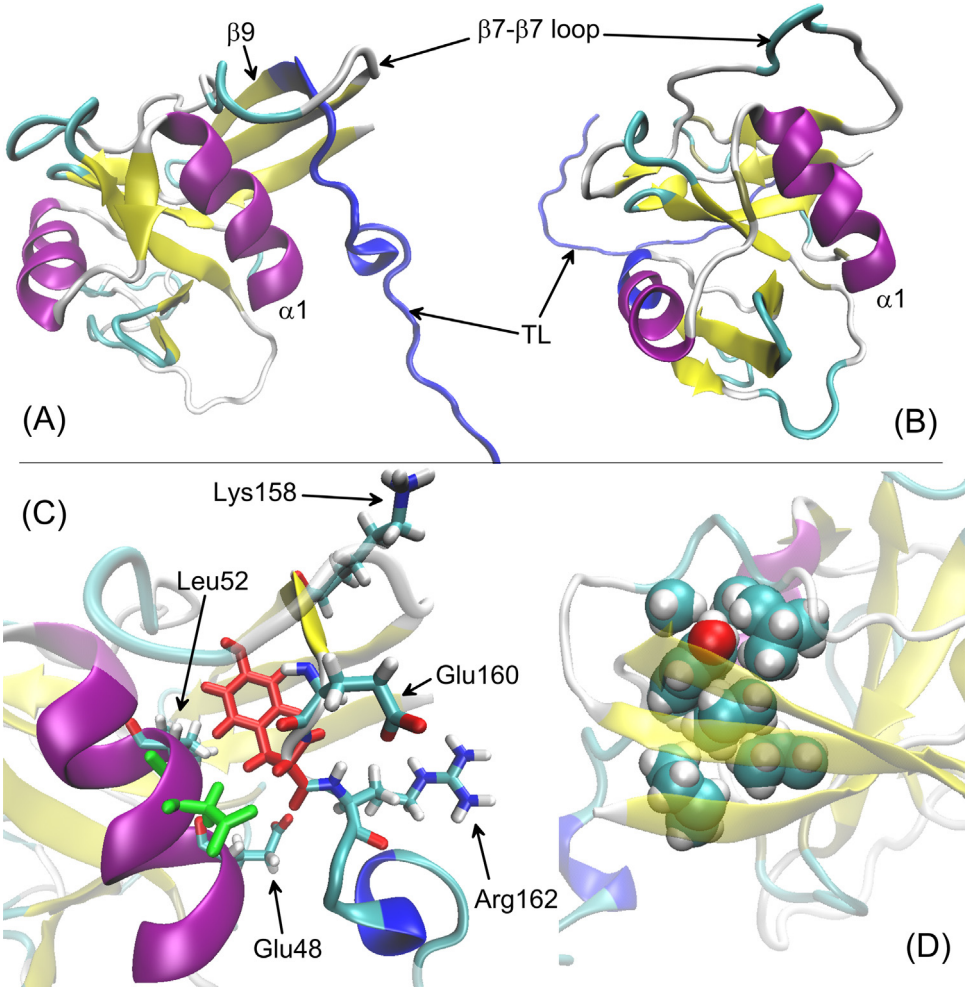


**Fig. 1.** (A) The 'ordered' (O) and (B) 'partially disordered' (PD) structures of the CD44 molecule, based on the PDB data (entries 1POZ and 2I83, respectively). The terminal lobe (TL) residues are marked in blue. (C) The vicinity of the Tyr161 residues showing some crucial amino acids (e.g. Glu48, Leu52, Lys158, Glu160 and Arg162). The residues defining the reaction coordinate (e.g. Tyr161 and Asp51) are shown in red and green, respectively. (D) Hydrophobic cluster found in the vicinity of Tyr161. Sidechains of Tyr161, Ile145, Ala138, Leu135, Ala55, Leu24 and Ile22 are shown in the ball representation. The pictures were prepared by using the VMD software [17].

force field is 'better' for reproduction of the HA-CD44 interactions, without further data. Moreover, it is worth noting that even the presence of a large free energy barrier separating the 'A' and 'B'-like states does not automatically mean that these two conformational forms exhibit different affinities for HA. The additional study is then required to be performed according to the procedures applied by Jamison et al. for obtaining the free energy changes accompanying the HA binding to CD44 in both 'A' and 'B' states (The convention of amino acid residues numbering in the above part of the paper corresponds to the murine CD44 and the X-ray structures by Banerji et al. [4]; in the further parts the numbering corresponding to the NMR structures (i.e. PDB: 1POZ and 2I83) will be used).

According to the NMR studies by Takeda et al. [2] and Ogino et al. [1], there exist two fundamental conformational forms of CD44; the first is very close to the structures obtained by Banerji et al., whereas in the other rearrangement of the β-strands in the extended lobe and disorder of the structure in the following C-terminal region can be observed. This observation and the subsequent studies led to the hypothesis that the order-to-disorder transition of the C-terminal region is induced by HA binding. However, it should be emphasized that both the 'ordered' (O) and 'partially disordered' (PD) forms of CD44 contribute to HA binding. Moreover both of them can exist in parallel either in the presence or absence of HA; the ligand appearance shifts the dynamic equilibrium state toward the PD form.

Fig. 1 presents schematically the structures representing the PD and O states of CD44 molecule. The range of the 'order-to-disorder' (O-PD) transition include mainly the lobular extension of the CD44 HA binding domain (HABD). The formation of the β-strand in the region corresponding to β9 of the O state is not observed in the PD one. In addition, β8 became rearranged relative to β0. Further, there is not enough data to state if the terminal residues (158–178) go under the loop between β7 and β8 as it is observed in the unbound (O) state. The residues 158–178 will be referred to as TL (terminal lobe). The detailed similarities and differences between PD and O forms are not reported here; the reader is referred to the original papers or to the related PDB entries. It is worth noting that the relation between the 'A'/'B' and O/PD pairs of conformational states has not been studied so far; the main problem is the lack information about the location of the HA ligand bound to the PD form of CD44 (the definitions of 'A' and 'B' are based on the Arg45-ligand distance).

The Tyr161Ala mutant exhibits both: (i) the PD conformation exclusively; (ii) much higher affinity for HA binding, compared to the native CD44. In the light of this correlation, it has been concluded that the PD form binds HA more strongly than the O one. However, understanding the reasons for such correlation is not trivial, as the Tyr161 residue lies nearly on the opposite side of the CD44 molecule with respect to the identified ligand anchoring site and the NMR structures of Tyr161Ala mutant do not exhibit any

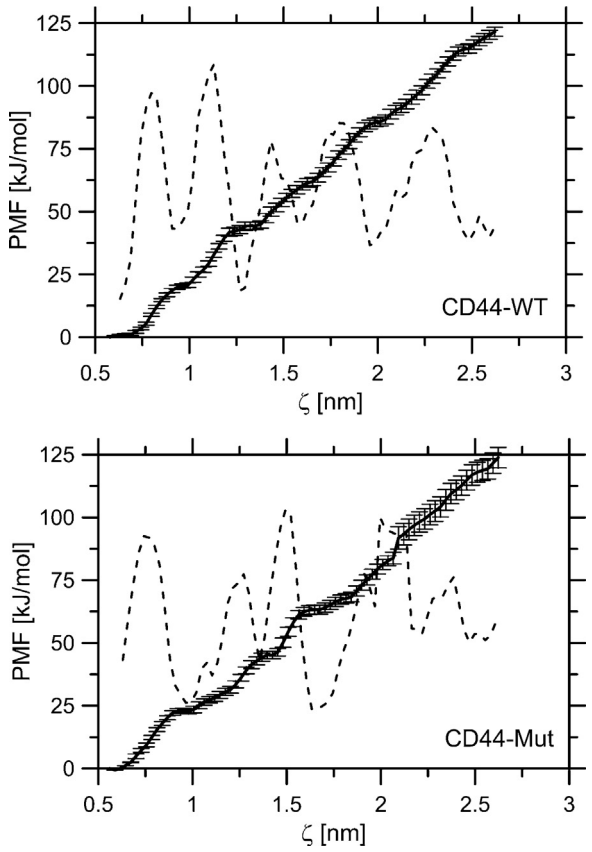


**Fig. 2.** The potential of mean force (PMF, solid lines) plots obtained for the initial stages of the O-PD transition for CD44-WT (top) and CD44-Mut (bottom). Vertical bars denote the errors obtained by the bootstrapping method. The PMF derivatives (dashed lines) are plotted to reflect the differences between PMFs better.

dramatic reorientation located in the direct vicinity of this binding cavity. Further investigation is hindered by the fact that, contrary to the O-like structures obtained in the XRD study [4], the position of the ligand (HA) bound to the PD conformational form of CD44 remains unknown.

Herein, we report the results of the molecular modeling study, focused on the possible mechanisms of the O-PD transition. Our aim was to: (i) identify the particular (initial) steps of the O-PD transition in the CD44 molecule; (ii) estimate the free energy profile associated with these steps; (iii) clarify the role of the selected amino acid residues (e.g. Tyr161) in the process of the O-PD transition. Both the wild-type CD44 and the Tyr161Ala mutant were subjected to the same research procedure to find out if their experimentally determined diverse behaviors will be theoretically confirmed.

As the extremely high free energy barriers are expected to accompany the O-PD transition (note, for instance that the rate constant for this process is of the order of ms), the routine molecular modeling techniques (e.g. unbiased molecular dynamics (MD)) are of the limited use. Here the steered molecular dynamics (umbrella sampling) method is applied to recover the free energy profile along the coordinate reflecting the studied conformational change. Furthermore, only the initial steps of the O-PD transition could be investigated; the reasons for that are explained in the Section 3.

Note that the above mentioned study is focused on the scenario according to which the PD structure resembles those deposited in the PDB database (entry: 2I83) rather than the (undefined) structure in which the TL residues go under the loop between β7 and β8 (e.g. 1POZ). Thus, the additional benefit from the reported study may be the identification of the more probable alternative.

## 2. Methods

The CD44 residues numbering is in accordance with that presented in ref. [1] and corresponds to the NMR structures (PDB: 2I83 and 1POZ).

### 2.1. MD simulations

The simulations were carried out using the CHARMM27 force field [6] as implemented [7] in the GROMACS 4.53 package [8]. The simulated systems included one protein molecule (wild type CD44; the initial structure taken from the PDB 1POZ entry or the manually prepared Tyr161Ala mutant) immersed in the simulation box containing water molecules (TIP3P model [9]) and 8 sodium ions to neutralize the total charge. Initially, the system was subjected to energy minimization on the whole system followed by the series of MD simulations with positional restraints: (i) immobile ('frozen') CD44 molecule; 100 ps; (ii) positional restraints on the protein backbone; the force constant equal to 10,000 kJ mol$^{-1}$ nm$^{-2}$; 2 ns; (iii) the same as in (ii) but the force constant equal to 1000 kJ mol$^{-1}$ nm$^{-2}$; (iv) the same as in (ii) but the force constant equal to 200 kJ mol$^{-1}$ nm$^{-2}$. Finally, the unrestrained MD simulations were performed for 50 ns (after 2.5 ns equilibration) and coordinates were saved to file every 2 ps for the subsequent analysis. The final structures were used in the subsequent umbrella sampling simulations. Additionally, the MD simulations were performed for the PD structure (2nd structure from the PDB entry 2I83 was accepted as the initial one) for duration varying from 5 to 60 ns, according to the protocol described above. The equations of motion were integrated with a timestep of 2 fs. The P-LINCS algorithm was applied to constrain all bond lengths [10]. The simulations were carried out under periodic boundary conditions based on the rectangular computational boxes (initial dimensions: 8.3 nm × 8.3 nm × 8.3 nm for the O-like structure, 9 nm × 9 nm × 9 nm for the PD-like structure and ~13 nm × 6.2 nm × 6.2 nm for the Umbrella Sampling simulations) containing ~23,300 (O), ~29,000 (PD) or ~15,700 (umbrella sampling) water molecules. The temperature was maintained close to its reference value (310 K) by applying the V-rescale thermostat [11] whereas for the constant pressure (1 atm, isotropic coordinate scaling) the Parrinello–Rahman barostat was used with a relaxation time of 1 ps [12]. The center of mass motion was removed every step. Electrostatics was treated with particle-mesh Ewald (PME) [13], using a short-range cutoff of 1.2 nm, and van der Waals interactions were switched off between 1.0 and 1.2 nm.

### 2.2. Umbrella sampling simulations

The umbrella sampling (US) approach obtains the free-energy profiles along a predefined coordinate from a set of equilibrated simulations. The US simulations were initiated from the O forms of the CD44 wild type (CD44-WT) and the CD44 Tyr161Ala mutant (CD44-Mut). The CD44 molecule was oriented to make the axis defined by the centers of masses (COMs) of Tyr161 (or Ala161 for CD44-Mut) and Asp51 approximately parallel to the Z-axis. The COM distance between the Tyr161/Ala161 and Asp51 residues was defined as a reaction coordinate $\xi$. The pull code in GROMACS [8] was used to generate from an initial pulling trajectory snapshots for the umbrella sampling simulations. Position restraints were imposed on the backbone atoms of α helices and β strands not belonging to TL and the Asp51 residue was used as an immobile reference for pulling simulations. The Tyr161/Ala161 residue was pulled away from Asp51 (together with the part of TL) along the Z-axis with a harmonic force constant of 5000 kJ mol$^{-1}$ nm$^{-2}$ and a pull rate of 0.01 nm ps$^{-1}$. A final COM distance of approximately 2.6 nm between Tyr161 and Asp51 was achieved. Along

the reaction coordinate, 32 windows were selected in the range $0.5\,nm \leq \xi \leq 2.6\,nm$ with a distance of ~0.06 nm between the adjacent positions. The data within each window were collected every 2 ps. The overall simulation time for a single PMF profile was about 1.28 μs (40 ns for each window). After removing the first 5 ns for equilibration, the PMFs were constructed with the weighted histogram analysis method (WHAM) [14] as implemented in GROMACS (*g_wham*) [8]. Statistical uncertainties of the PMFs were estimated using the Bayesian bootstrapping of complete histograms [15]. The comments on the choice of the given reaction coordinate are given in the subsequent section.

## 3. Results and discussion

Observing the process of the full O-PD transition is difficult even if applying the steered molecular dynamics protocols. The unbiased MD simulations of the PD structure revealed that after several nanoseconds the TL chain begins to attach to the remaining part of the HABD domain. Furthermore, no preferential interactions involving Tyr161 were observed. The resulting structures differ from each other, depending on the simulation run but the spontaneous formation of the O-like structure was not observed. Further propagation of the simulation does not change the structure much, indicating that the system has been trapped in the deep (different) local minima of the free energy. The entire collection of the final structures resulting from this stage of the study resembles closely that of the NMR structures deposited under the PDB 2I83 record. The only (qualitative) difference is that the MD-resulting structures exhibit a slightly larger area of the contact of the TL residues with the remaining part of the CD44 molecule (this, however, can be ascribed to the minor inaccuracies in the force field or to the limited time of sampling). Thus, even if designing the perfect reaction coordinate reflecting the full O-PD transition process, the two following hindrances will appear: (i) the problem of definition of the PD state: both the NMR structures and MD simulations indicate that the PD form is rather a collection of different structures differing significantly in the location of their TLs (this is rather understandable as TL is the disordered part of the molecule). (ii) The above mentioned structures are separated by high free energy barriers which could not be overcome during the unbiased MD simulation lasting tens of nanoseconds. Thus, no efficient sampling can be expected for the collection of the PD-like states during possible US simulation. The potential solution might be the application of the multidimensional US-related techniques but the large dimensions of the system throw doubts on the effectiveness of such methods.

In view of the above difficulties, we decided to study only the initial steps of the O-PD transition, which seem to be the crucial ones for the overall PD/O equilibrium, due to the experimentally confirmed role of Tyr161.

The O-like structures were accepted for both CD44-WT and CD44-Mut which may seem unrealistic in the latter case as the Tyr161Ala adopts the PD conformation. However, if (hypothetically) assuming that the differences caused by the mutation results exclusively from the action of Tyr161 (or Ala161) in the region limited by the loop between β7 and β8 (i.e. the vicinity of Tyr161 in the O-like structure) then one can expect the significant differences in the corresponding PMF curves. This makes the US investigation focused on the O-like initial form of CD44-Mut valuable. Furthermore, the process of dynamic equilibrium between PD- and O-like structures observed for CD44-WT let to speculate that the similar type of equilibrium may also exist in the case of CD44-Mut but, due to the favorableness of the PD form, is difficult for being experimentally monitored.

First, we would like to note that the standard (unbiased) MD simulations did not lead to any significant rearrangements in the TL region which could be ascribed to the O-PD transition or its initial stages neither for CD44-WT nor CD44-Mut. It is not surprising in the case of CD44-WT, as the very high energy barriers are expected to be connected with this process. It is also in agreement with the previous molecular modeling simulations [3]. The same observation for the CD44-Mut speaks for the conclusions that the O-PD transformation is the activated process in both cases. The additional contribution of this part of the study was the preparation of the fully equilibrated structures for the subsequent pulling simulations.

According to the molecular structures of the O conformation and to Ref. [1], the interactions of Tyr161 (TL) with Glu48 and Leu52 (α1 helix) seem to be critical for stabilizing the O conformation. For studying the O-PD transformation we have chosen the coordinate associated with the Asp51 residue, located between Glu48 and Leu52, being close to Tyr161 but not exhibiting any strong attractive interaction with it. The achieved range of the $\xi$ values corresponds to relative smooth transition from the O-like structure ($\xi \sim 0.6\,nm$) to the intermediate structure ($\xi \sim 2\,nm$) in which the Arg162 sidechain is outside the region limited by the loop between β7 and β8 and interacts directly with Asp140.

The free energy profiles (PMF curves) obtained for CD44-WT and CD44-Mut are presented in Fig. 2. The most striking observation is their similarity: more significant differences are visible only after calculating the PMF derivative after $\xi$ (i.e. the force curve). Even after that, the four large (>75 kJ/mol/nm) force maxima can be distinguished in both cases. No minima or maxima can be observed on the PMF curves, indicating the lack of local minima or maxima of the free energy associated with the $\zeta$ coordinate. The global minimum of the free energy can be ascribed to the initial (O-like) structure for both CD44-WT and CD44-Mut. The process of TL unfolding is connected with the monotonous increase of the PMF function values. It is worth noting at this stage of results analysis that these data neither confirm nor negate the hypothesis for the dynamic equilibrium between the PD and O conformation and on its shift toward PD in the CD44-Mut molecule. This type of equilibrium should be referred to the final PD structure which cannot be sampled in the presented US study (see comments above). In other words, we do not know if the PD form corresponds to the minimum of the free energy and what is the difference between this (hypothetical) minimum and the global minimum obtained for the O conformation. For the same reasons, we can not state if the PD/O equilibrium is shifted toward the PD form for the case of CD44-Mut.

Below some characteristic regions of the PMF plots are briefly described. We start with the molecular details of the interactions resulting in the appearance of the global minimum corresponding to the O-like conformation of the CD44-WT molecule. The common attractive interactions involving the TL residues include hydrogen bonding (HB) between the following residue pairs: Glu48-Arg162, Phe139-Glu160, Leu52-Tyr161, Glu48-Tyr161. Furthermore, HB is observed within the TL part itself for the Glu160-Tyr161 and Glu160-Thr163 pair. Additionally, the appearance of ionic bridges can be observed for the Glu48-Arg162 (independently of HB listed above) and Asp175-Lys158. The latter interaction is especially interesting as it involves one of the terminal residues (Asp175) and Lys158 located on the other side of the β7-β8 loop, confining the TL part. Thus, the additional, closed 'loop' is formed in the distance from Lys158 to Asp175 with the β7-β8 loop placed inside. This effect obviously limits the mobility of TL and seems to promote the stability of the O conformation. However, similar interaction was not observed for the case of CD44-Mut and the related PMF plot does not reflect any significant difference (see the description below). It can also be supposed that the presence or the absence of this ionic bridge is connected with the large conformational space accessible for the mobile ending of TL and not all relevant conformations were sampled during 50 ns of the simulations. Apart from the above listed interactions, there can be noted a contribution of

the weak hydrophobic interactions involving sidechains of numerous residues belonging to TL, the α1 helix or β-strands at the same time. In other words, the hydrophobic cluster is formed with the contribution of the following residues: Tyr161 (the hydroxyl group is exposed out of the cluster), Ile145, Ala138, Leu135, Ala55, Leu24 and Ile22.

The analyzed PMF does not contain any characteristic extremes, thus, the attention will be paid to the minima of the force plot (i.e. PMF derivative) which are the closest to the favorable intramolecular interactions. The differences between the first local minimum of force (located around ~0.9 nm) and the global minimum of PMF can be ascribed to the following molecular processes: (i) disrupting most of the HBs between TL and the rest of the protein and breaking the Asp175-Lys158 ionic bridge; (ii) the increased distance between TL and α1 and the increase of the TL mobility; (iii) formation of the new, attractive interactions, such as the Asp51-Arg162, Glu160-Arg162 and Lys158-Asp140 ionic bridges and Lys150-Pro142 HB (competitive with respect to the latter ionic bridge); (iv) the hydrophobic cluster in the vicinity of Tyr161 is maintained.

The second local minimum of the force (~1.3 nm) is accompanied by the breaking of the above listed ionic bridges and maintaining only the Lys158-Pro142, Tyr161-Ala55 and Arg162-Asp51 (without the contribution of Asp51 sidechain, thus, it is not classified as the ionic bridge) hydrogen bonds. The two latter HBs are the only direct attractive interactions between TL and α1. Also the hydrophobic cluster undergoes collapse. Note that at this point, the main molecular processes occur in the vicinity of the TL/α1 edge of contact and the location of the TL buried under the β7-β8 loop remains unchanged.

This situation is changed when considering two further minima of the force plot. For $\zeta > 1.5$ nm, the most intensive interactions involving the TL residues are concentrated around the region of the TL contact with the β7-β8 loop. Around $\zeta = 1.6$ nm the TL part is pulled outside the β7-β8 loop up the Glu160 residue. Thus, the maximum of the force around ~1.5 nm can be ascribed to overcoming the steric hindrances which accompany the translocation of the two large Tyr161 and Glu160 sidechains under the limiting β7-β8 loop. Furthermore, the stable, intra-TL HB is developed between Arg162 and Asn164.

The final (i.e. the last analyzed) minimum around $\zeta = 2$ nm corresponds to the structure in which the TL residues are pulled out from the β7-β8 loop further, up to the Arg162 residue, which (being outside the region limited by the loop) forms the ionic bridge with Asp140. Additionally, Tyr161 creates HB with Gly141. In this stage of the O-PD transition, the interactions between TL and the rest of the CD44 molecule are concentrated around the β7-β8 loop region.

The additional description for the case of CD44-Mut is much more condensed due to the following fundamental fact: nearly all the characteristic intramolecular interactions observed for CD44-WT can be also monitored for CD44-Mut. This obviously does not concern the interactions involving Tyr161, replaced by Ala161 in CD44-Mut, but the differences resulting from such replacement are surprisingly small (the example of such a difference may be the shallower and shifted fourth local minimum of the force plot obtained for the CD44-Mut case; this is correlated with the absence of the Tyr161-Gly141 hydrogen bond). For instance, the small sidechain of Ala161 is the part of the hydrophobic cluster which appears also for the CD44-Mut. Furthermore, one may notice slight changes in the characteristic values of $\zeta$, describing the position of the local minima on the force curve; this effect can be ascribed to the differences in the definitions of $\zeta$ coordinate, holding for CD44-WT and CD44-Mut (distances Tyr161-Asp51 vs. Ala161-Asp51). The most notable examples are shifts of the second and fourth peaks of the force curve corresponding to CD44-Mut when comparing with the peaks obtained for CD44-WT. No other significant qualitative differences between the molecular mechanisms of the O-PD transitions studied for both CD44-WT and CD44-Mut were monitored.

There is no clear explanation for the observations described above. Below we discuss several potential reasons for the likeness of the two analyzed PMF plots.

There exists a possibility that the role of the Tyr161 residue becomes important in the further stages of the O-PD transition (e.g. Tyr161, contrary to Ala161, could prevent from such specific interaction stabilizing PD conformation). Due to the above mentioned sampling problems, verifying such hypothesis is not possible in the current state of knowledge and due to the limited computational resources.

The competitive explanation requires arguing against the PD-like structures in which the TL part is located outside the limiting β7-β8 loop. The competitive PD-like structures could exhibit greater resemblance to the O-like structure in which the TL chain goes under the β7-β8 loop and the lack of interactions between TL and α1 is caused by other (unidentified yet) conformational rearrangements in the TL region resulting in its unfolding. Such assumption was considered in Ref. [1] but no experimental verification was possible.

Independently of the correctness of any of the above explanations, there are several important conclusions which can be drawn on the basis of the reported study. First of them is that the differences in the behaviors exhibited by CD44-WT and CD44-Mut do not simply result from the strength of interactions between the Tyr161-Glu48 and Tyr161-Leu52 pairs. The disruption of these interactions probably contributes to the overall PMF increment but this contribution is rather slight as in the case of the CD44-Mut, for which the above interaction does not exist, the PMF course is very similar. Secondly, the overall HB network and another type of interactions projected along the PMF curve have rather conservative character and seem to be independent of the Tyr161Ala mutation; the exception are the HB-type interactions involving the hydroxyphenyl moiety of Tyr161 which result in only minor differences between PMFs. This includes the global minima corresponding to both studied structures.

Let us focus on the comparison of the experimentally estimated time scale of the O-PD conformational transition (hundreds of milliseconds, according to Ref. [1]) and the theoretically calculated free energy barrier. In the absence of both the full free energy profile linking the O and PD forms and the kinetic information about the transition trajectories, there is no possibility of direct comparison of measurable quantities. However, according to the transition state theory [16] the free energy barrier height corresponding to such experimental estimation of the forward reaction rate constant (O → PD) (with the transmission coefficient set to unity) as of the larger order than the ~100 kJ/mol, i.e. the PMF value for the maximum $\zeta$. The height of 100 kJ/mol correspond to the reaction rate around $9 \times 10^{-5}\ s^{-1}$. Moreover, it is very probable that the free energy landscape for further progress of the O-PD transition is more complicated than its initial stages analyzed here making further studies more complicated. Nevertheless, the main conclusion is that the large values of the calculated PMF are not contradictory to the experimental data. The competitive explanation is that the large free energy values may be the result of the lack of convergence. Although the WHAM input histograms overlap each other (which is an essential condition), one cannot exclude the possibility that the system is trapped in the local but deep free energy minimum somewhere along the O-PD reaction path. A similar situation has been considered during discussion on the results of the unbiased MD simulations.

Let us finally note that the above presented qualitative analysis of the intramolecular interactions does not include the full description of the dynamic properties of the system (e.g. the HB

lifetimes) as the US approach influences such properties leaving only the thermodynamic properties to be calculated quantitatively.

## 4. Conclusions

The molecular dynamics/umbrella sampling simulations of the wild-type CD44 and the Tyr161Ala CD44 mutant allowed for estimation of the free energy profiles associated with the initial stages of the 'order-to-disorder' (O-PD) conformational transition postulated to occur in the terminal lobe of the CD44 molecule. This let for the identification of the most important interactions in the vicinity of the Tyr161 residue which is believed to be crucial for the process of the O-PD transition. The results are somehow surprising and can be summarized that both the wild-type CD44 and the Tyr161Ala CD44 mutant exhibit roughly the same free energy profiles accompanying the O-PD rearrangement. This led to the conclusions that the role of Tyr161 is not simply limited to its attractive interaction involving the Glu48 and Leu52 residues. The function of Tyr161 is not definitely verified and requires further studies. At present the two competitive hypotheses can be proposed aiming at the explanation of the close resemblance of the free energy profiles corresponding to the two studied molecules. The first of them assumes the existence of some (unknown) interactions located outside the β7-β8 loop which are stabilized by Ala161 but destabilized by Tyr161. The second one requires accepting the PD-like structure in which the terminal lobe remains located under the β7-β8 loop (i.e. unlike the NMR-based PD-like structures).

## References

[1] S. Ogino, N. Nishida, R. Umemoto, M. Suzuki, M. Takeda, H. Terasawa, J. Kitayama, M. Matsumoto, H. Hayasaka, M. Miyasaka, I. Shimada, Two-state conformations in the hyaluronan-binding domain regulate CD44 adhesiveness under flow condition, Structure 18 (2010) 649–656.
[2] M. Takeda, S. Ogino, R. Umemoto, M. Sakakura, M. Kajiwara, K.N. Sugahara, H. Hayasaka, M. Miyasaka, H. Terasawa, I. Shimada, Ligand-induced structural changes of the CD44 hyaluronan-binding domain revealed by NMR, J. Biol. Chem. 281 (2006) 40089–40095.
[3] W. Plazinski, A. Knys-Dzieciuch, Interactions between CD44 protein and hyaluronan: insights from the computational study, Mol. BioSyst. 8 (2012) 543–547.
[4] S. Banerji, A.J. Wright, M. Noble, D.J. Mahoney, I.D. Campbell, A.J. Day, D.G. Jackson, Structures of the Cd44-hyaluronan complex provide insight into a fundamental carbohydrate-protein interaction, Nat. Struct. Mol. Biol. 14 (2007) 234–239.
[5] F.W. Jamison 2nd, T.J. Foster, J.A. Barker, R.D. Hills Jr., O. Guvench, Mechanism of binding site conformational switching in the CD44-hyaluronan protein-carbohydrate binding interaction, J. Mol. Biol. 406 (2011) 631–647.
[6] A.D. Mackerell Jr., M. Feig, C.L. Brooks 3rd, Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, J. Comput. Chem. 25 (2004) 1400–1415.
[7] P. Bjelkmar, P. Larsson, M.A. Cuendet, B. Hess, E. Lindahl, Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models, J. Chem. Theory Comput. 6 (2010) 459–466.
[8] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718.
[9] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[10] B. Hess, P-LINCS: a parallel linear constraint solver for molecular simulation, J. Chem. Theory Comput. 4 (2008) 116–122.
[11] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity-rescaling, J. Chem. Phys. 126 (2007) 014101.
[12] M. Parinello, A. Rahman, Polymorphic transitions in single crystals: a new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190.
[13] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8593.
[14] S. Kumar, D. Bouzida, R.H. Swendsen, P.A. Kollman, J.M. Rosenberg, The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method, J. Comput. Chem. 13 (1992) 1011–1021.
[15] J.S. Hub, B.L. de Groot, D. van der Spoel, g_wham – a free weighted histogram analysis implementation including robust error and autocorrelation estimates, J. Chem. Theory Comput. 6 (2010) 3713–3720.
[16] H. Eyring, The theory of absolute reaction rates, Trans. Faraday Soc. 34 (1938) 41–48.
[17] W. Humphrey, A. Dalke, K. Schulten, Vmd – visual molecular dynamics, J. Mol. Graphics 14 (1996) 33–38.