Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Bioorganic & Medicinal Chemistry Letters 14 (2004) 1765–1769
Bioorganic & Medicinal Chemistry Letters

Protein-dynamics of the putative HCV receptor CD81 large extracellular loop☆

Alexander Neugebauer, Christian D. P. Klein and Rolf W. Hartmann*

Pharmaceutical and Medicinal Chemistry, Saarland University, PO Box 151150, D-66041 Saarbrücken, Germany

Received 29 October 2003; revised 8 January 2004; accepted 14 January 2004

Abstract—The human CD81 protein is a likely receptor for the binding of hepatitis C virus (HCV) to hepatocytes and therefore a possible target for novel anti-HCV drugs. The two published X-ray structures of the HCV binding region of CD81 (1G8Q and 1IV5) have, particularly in a substructure that is formed by two helices, a slightly different conformation. The abovementioned substructure is a candidate target region for virtual screening approaches. We present here a molecular dynamics study of the two X-ray structures. Our results indicate that the conformation of the two helical regions in one of the X-ray structures (1G8Q) is affected by crystallographic contacts and most likely does not represent the native state of the protein.

1. Introduction

Hepatitis C virus (HCV), a positive-stranded RNA virus, is the infectious agent responsible for one subtype of viral hepatitis. Today, 300 million people worldwide are infected by HCV. The therapy of choice currently consists of a combination of ribavirin and (pegylated) interferon, but this therapeutic regimen is expensive and fails in about 40% of all cases.[1] Novel therapeutical strategies are therefore urgently needed.

The HCV genome encodes for a single polyprotein of approximately 3000 amino acids. The polyprotein is processed by host and viral peptidases into different viral proteins.[2] Among these proteins, the envelope glycoprotein E2 is supposed to be the key protein for viral entry.

Putative receptors for HCV–E2 binding interaction are the human CD81 receptor,[3–5] the low-density lipoprotein receptor (LDLR),[6] the human scavenger receptor class B type I[7] and liver/lymph node-specific intercellular adhesion molecule-3-grabbing integrin (L-SIGN).[8]

Among these receptors, X-ray structures are available for the large extracellular loop (LEL) of CD81 only. Furthermore, the interaction sites with the E2 protein have been identified recently.[3,9] This makes the LEL on CD81 a potential drug target. Recent results by Van-Compernolle[10] and by our own group (unpublished data) indicate that small molecules which interfere with the CD81/E2 interaction can inhibit the binding of HCV to their host cells.

The three-dimensional structure of LEL was determined by Kitadokoro et al. using X-ray crystallography and two different crystallographic forms (monoclinic: pdb ID 1G8Q and hexagonal: pdb ID 1IV5).[11,12] An important difference between the two structures is a cleft-like motif formed by the C and D helices in the 1G8Q structure which does not appear in the 1IV5 structure. Therefore, different classes of inhibitors for the E2–CD81 interaction process seem to be possible: In the case of the 1G8Q structure, small organic molecules fit into the abovementioned cleft, whereas in the 1IV5 form only surface interactions between the protein and the ligand are possible. Here, we investigate the dynamics of the LEL region of the CD81 receptor by molecular dynamics simulations to gain deeper insight in the CD81–LEL drug target. Our main objective was to

Keywords: Hepatitis C virus; CD81 large extracellular loop; HCV E2 glycoprotein; Molecular dynamics simulations; Protein–protein interactions.
☆Supplementary data associated with this article can be found, in the online version at, doi: 10.1016/j.bmcl.2004.01.036
* Corresponding author. Tel.: +49-681-302-2424; fax: +49-681-302-4386; e-mail: rwh@mx.uni-saarland.de

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.01.036

determine the physiologically relevant conformation of the LEL region, which can be used as the target protein in virtual high-throughput screening.

2. Methods

X-ray structures of the monoclinic (pdb ID: 1G8Q) and hexagonal (pdb ID: 1IV5) crystallization form of the CD81-LEL were obtained from the Brookhaven Protein Databank (PDB). All molecular dynamic simulations were performed by the AMBER 7 suite of programs.[13] Water molecules were removed from the PDB files and missing protein atoms were added. Disulfide bond formation and charge neutralization (addition of Na^{+-} ions) were carried out using the AMBER utility Xleap. Afterwards, the simulation systems were minimized (5000 steps of steepest descent followed by 10,000 steps of conjugate gradient) and equilibrated at 300 K (20,000 steps, step size 1 fs). Finally, in both cases 1IV5 and 1G8Q 1500 ps of molecular dynamics simulation (step size 1.5 fs, 1,000,000 steps) at constant temperature (300 K) were performed. Temperature regulation was done by coupling to an external bath (Berendsen's method: see ref 14) using a bath coupling constant of 1.0 ps during equilibration and 1.5 ps during production. The generalized Born model was used to account for solvation effects.[15]

Crystallographic symmetries were built using the Swiss-PdbViewer and the symmetry related information within the original PDB-files of the CD81–LEL structures. Superpositions of molecular structures were also performed by Swiss-PdbViewer.[16]

3. Results and discussion

3.1. Molecular dynamics simulations

CD81–LEL was crystallized as dimer. Each monomer consists of five helices (A–E) where helices A and B provide the main dimer association interface. The E2 binding region is located close to the C and D helices. Therefore, the structures of the C and D helices are important for E2 and inhibitor binding. To evaluate their dynamic behavior, we calculated the average mass-weighted fluctuations of all amino acids of both LEL structures. The calculation was performed on the last 400 ps of the MD trajectory. These kind of fluctuations are only valid when the protein system is in equilibrium. Therefore Figure 1 shows the energy versus time plots of the 1G8Q and 1IV5 MD simulations. Figure 2 shows the average mass-weighted fluctuation of all atoms of all residues plotted against the residue numbers. The dashed line corresponds to the 1G8Q and the solid line to the 1IV5 structure. The fluctuation of the 1IV5 protein is much lower than the fluctuation of the 1G8Q structure. The average fluctuation of all residues is 1.379 Å for 1IV5 and 3.403 Å for 1G8Q.

The flexibility of the helices C and D is somewhat larger than the flexibility of the helices A, B and E (Fig. 2). This is valid for both crystallographic forms. The main reason for the lower fluctuations of the A and B helices is the formation of the main dimer association interface between the LEL monomers. In the 1G8Q structure, pronounced fluctuations are observed in the C and D helices. Six amino acids in these two helices (Asp137, Asp138, Asp139, Ile181, Ile182, Ser183) show average mass-weighted fluctuations of more than 5 Å. For comparison, no residue in the 1IV5 structure has such a high degree of mobility.

The same procedure was repeated for the backbone atoms to estimate the influence of backbone motions. The average backbone atom fluctuations are 1.324 Å (1IV5) and 3.253 Å (1G8Q), respectively. This indicates that the average all-atom fluctuations result mainly from the backbone movement.

The interaction between the amino acids of the C and D helices are responsible for the close-down of the

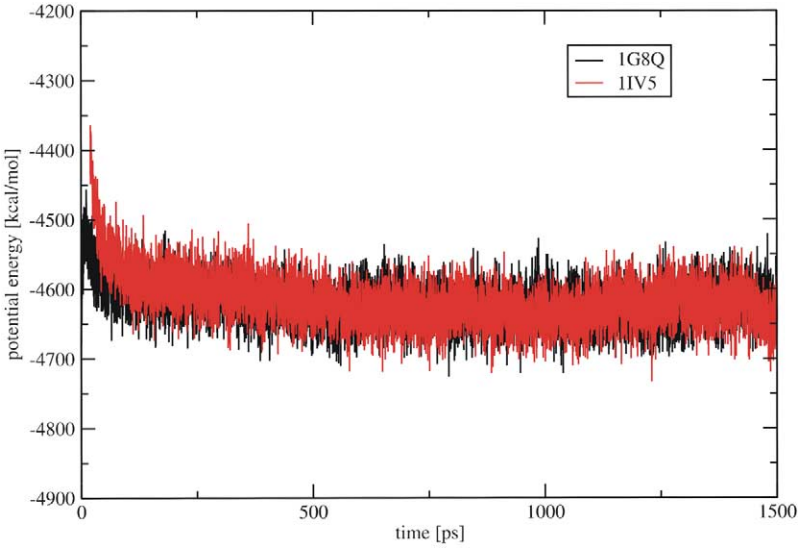

Figure 1. Energy versus time plots of the 1G8Q and 1IV5 protein systems.

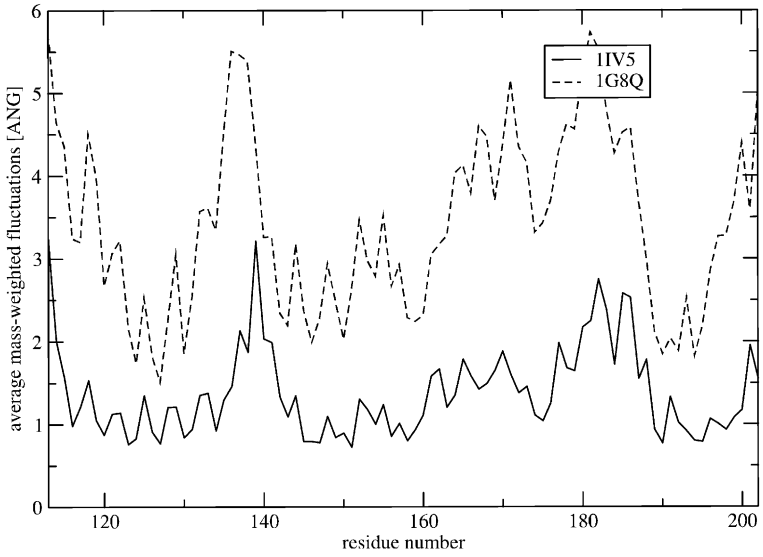

Figure 2. Dynamics of CD81-LEL: Average mass-weighted fluctuations of all residues during the last 400 ps of the MD simulation. Dashed line: 1G8Q, solid line: 1IV5. The letters below the curves indicate the helical regions A–E.

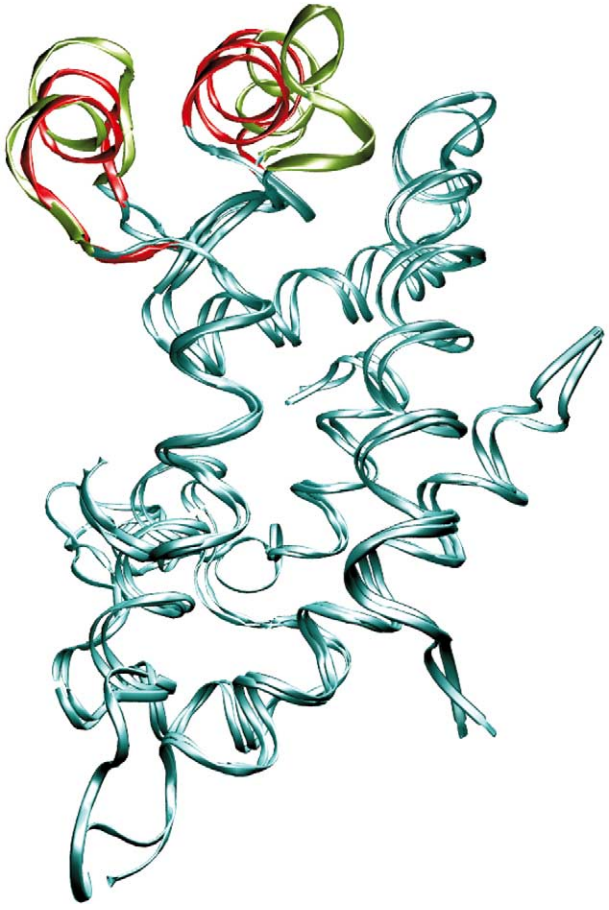

Figure 3. Superposition of the 1G8Q (green substructure) and 1IV5 (red substructure) CD81-LEL crystal structures.

CD-cleft after a simulation time of 50 ps. In particular, van der Waals and hydrogen bonding interactions occur between residues Leu185–Thr163, Asn184–Thr166/Thr167, Ile181–Thr167, Asn180–Leu170, Phe186–Thr163 and Ile181–Lys171.

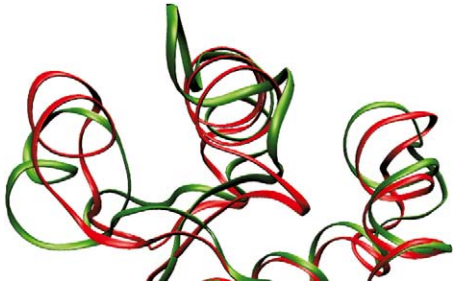

Figure 4. Superposition of the C and D helices of the 1IV5 cystal structure (red) and 1G8Q after 1.5 ns of molecular dynamics simulation (green).

This observation was verified by performing the 1G8Q MD simulation five times (100 ps production phase) using different (random) initial velocities. Each simulation leads to the closure of the CD cleft after approximately 50 ps. The structural deviations from the 1G8Q crystal structure for each simulation are 1.27, 1.30, 1.36, 1.27 and 1.27 Å. Each calculation leads to a comparable final structure.

We therefore conclude that the hexagonal 1IV5 structure is the predominating conformation under physiological conditions.

3.2. Ramachandran plots

In the monoclinic 1G8Q structure, more amino acids are outside the energetically favored regions of the Ramachandran plot (16 after 100 ps MD, 8 after 500 ps MD) than in the 1IV5 structure (2 after 100 ps MD, 4 after 500 ps MD). There is a reduction in amino acid numbers occupying unfavored areas of the Ramachandran plot for 1G8Q, but the 1IV5 values are not reached. This observation also supports the notion that the 1IV5 structure is preferred under physiological conditions.

3.3. Superposition of CD81–LEL structures

Figure 3 shows the superposition of the monoclinic 1G8Q (green) and hexagonal 1IV5 (red) crystal structures of the CD81–LEL region. The image also emphasizes the cleft between the C and D helices in the 1G8Q structure. The main difference is the position of the C helix which locks the cleft in the 1IV5 form. The D helix is in a comparable position in both structures.

Figure 4 shows the superposition of the C and D helices of the 1IV5 (red) crystal structure and 1G8Q (green) after 1.5 ns of molecular dynamics simulation. The root mean square (RMS) of the amino acids forming the cleft is 1.51 Å. This picture clearly shows that the 1G8Q structure approaches the geometry of the 1IV5 crystal structure during the MD simulation. The C and D helices in the 1G8Q structure after MD occupy the same region as compared to the 1IV5 crystal structure. The 1G8Q structure is not completely, but to a large extent, converted to the 1IV5 structure during the MD simulation.

3.4. Crystallographic contacts

In both crystallographic structures (1G8Q and 1IV5) crystallographic contacts are present for the C and D helices. However, these contacts are completely different. In the monoclinic 1G8Q structure, the D helix performs interactions with the C′ and D′ helices of an adjacent LEL molecule. Equivalent interactions occur

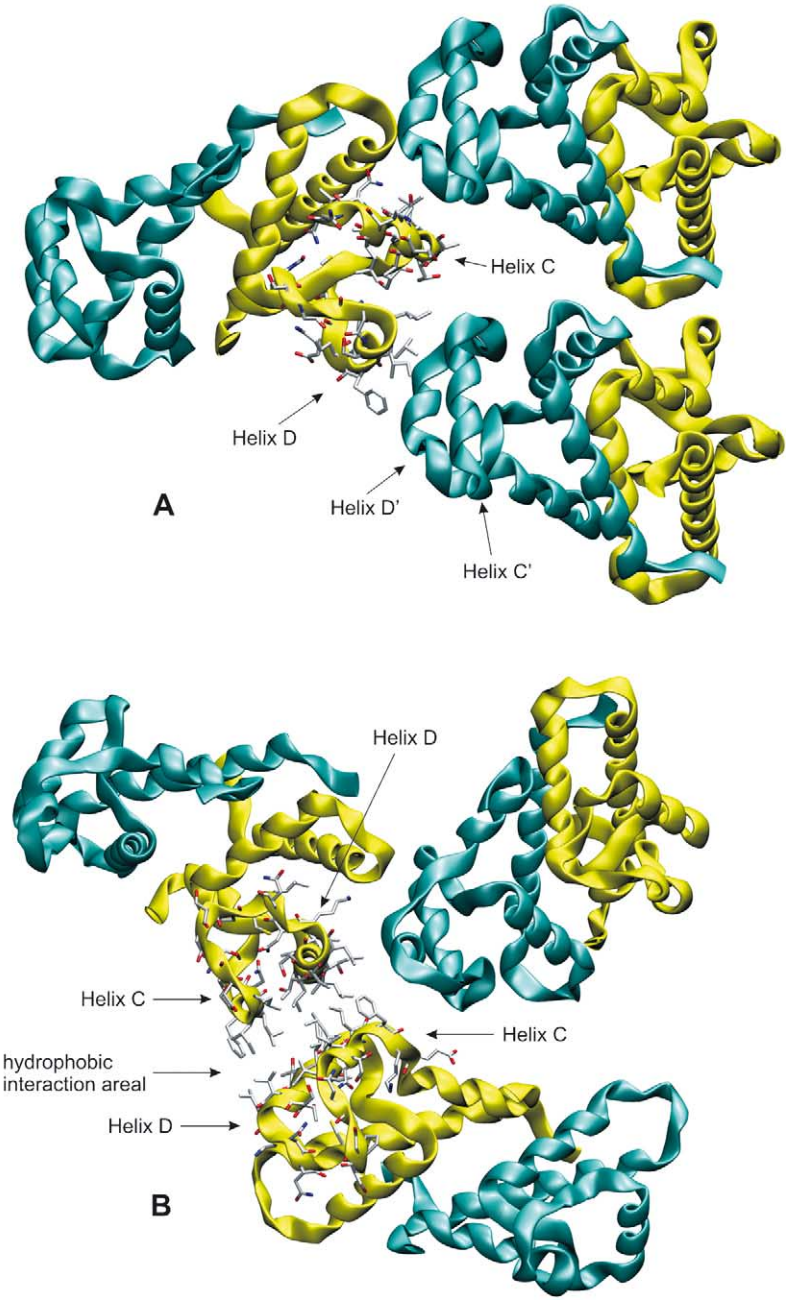

Figure 5. Crystallographic contacts of 1G8Q (**A**) and 1IV5 (**B**) crystal structures of the CD81–LEL. The yellow and the turquoise substructures denote the monomers of the LEL dimer.

to the C helix by a third LEL molecule. Figure 5A shows the crystallographic contacts of the monoclinic 1G8Q lattice.

In the hexagonal 1IV5 structure, the C and D helices form an α-helical bundle with the C and D helices of another LEL molecule (Fig. 5B). This arrangement should be more stable than the corresponding arrangement in the monoclinic 1G8Q case, because there are much more intensive crystallographic contacts. Furthermore, the contact area in the hexagonal 1IV5 structure with neighboring LEL molecules is strongly hydrophobic (see Fig. 5B). Because of the reasons given above, the cleft formation is suppressed in the hexagonal 1IV5 lattice.

4. Conclusions

The simulations have, combined with the analysis of the X-ray structures, shown that crystallographic contacts cause a distortion of the CD81–LEL folding in one of the two X-ray structures. Under physiological conditions, CD81–LEL most likely exists in the conformation with a closed cleft between the C and D helices. However, the two helices experience a certain degree of flexibility, and it may be possible to identify—for example, by virtual screening—small molecules that fit into the cleft and thereby prevent the interaction between CD81 and E2.

Acknowledgements

We appreciate financial support by the German 'Fonds der Chemischen Industrie' (FCI).

References and notes

1. Fried, M. W.; Shiffman, M. L.; Reddy, K. R.; Smith, C.; Marinos, G.; Goncales, F. L., Jr.; Haussinger, D.; Diago, M.; Carosi, G.; Dhumeaux, D.; Craxi, A.; Lin, A.; Hoffman, J.; Yu, J. *N. Engl. J. Med.* **2002**, *347*, 975.
2. Rosenberg, S. *J. Mol. Biol.* **2001**, *313*, 451.
3. Higginbottom, A.; Quinn, E. R.; Kuo, C. C.; Flint, M.; Wilson, L. H.; Bianchi, E.; Nicosia, A.; Monk, P. N.; McKeating, J. A.; Levy, S. *J. Virol.* **2000**, *74*, 3642.
4. Petracca, R.; Falugi, F.; Galli, G.; Norais, N.; Rosa, D.; Campagnoli, S.; Burgio, V.; Di Stasio, E.; Giardina, B.; Houghton, M.; Abrignani, S.; Grandi, G. *J. Virol.* **2000**, *74*, 4824.
5. Tan, Y. J.; Lim, S. P.; Ng, P.; Goh, P. Y.; Lim, S. G.; Tan, Y. H.; Hong, W. *Virology* **2003**, *308*, 250.
6. Germi, R.; Crance, J. M.; Garin, D.; Guimet, J.; Lortat-Jacob, H.; Ruigrok, R. W.; Zarski, J. P.; Drouet, E. *J. Med. Virol.* **2002**, *68*, 206.
7. Scarselli, E.; Ansuini, H.; Cerino, R.; Roccasecca, R. M.; Acali, S.; Filocamo, G.; Traboni, C.; Nicosia, A.; Cortese, R.; Vitelli, A. *Embo. J.* **2002**, *21*, 5017.
8. Gardner, J. P.; Durso, R. J.; Arrigale, R. R.; Donovan, G. P.; Maddon, P. J.; Dragic, T.; Olson, W. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 4498.
9. Drummer, H. E.; Wilson, K. A.; Poumbourios, P. *J. Virol.* **2002**, *76*, 11143.
10. VanCompernolle, S. E.; Wiznycia, A. V.; Rush, J. R.; Dhanasekaran, M.; Baures, P. W.; Todd, S. C. *Virology* **2003**, *314*, 371.
11. Kitadokoro, K.; Ponassi, M.; Galli, G.; Petracca, R.; Falugi, F.; Grandi, G.; Bolognesi, M. *Biol. Chem.* **2002**, *383*, 1447.
12. Kitadokoro, K.; Bordo, D.; Galli, G.; Petracca, R.; Falugi, F.; Abrignani, S.; Grandi, G.; Bolognesi, M. *Embo J.* **2001**, *20*, 12.
13. Case, D. A.; Pearlman, D. A.; Caldwell, J. W. T. E. C., III; Wang, J.; Ross, W. S.; Simmerling, C.; Darden, T.; Merz, K. M.; Stanton, R. V.; Cheng, A.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. *AMBER 7*; University of California: San Francisco, 2002.
14. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
15. Tsui, V.; Case, D. A. *Biopolymers* **2000**, *56*, 275.
16. Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714.