The euHCVdb Suite of *In Silico* Tools for Investigating the Structural Impact of Mutations in Hepatitis C Virus Proteins

C. Combet[#], E. Bettler[#], R. Terreux[#], N. Garnier and G. Deléage*

Institut de Biologie et Chimie des Protéines; UMR5086; CNRS; Université Lyon 1; IFR128 BioSciences Lyon-Gerland; 7, passage du Vercors, 69367 Lyon CEDEX 07, France

Abstract: Hepatitis C is a viral infection of the liver that results in acute hepatitis and chronic liver disease, including cirrhosis and liver cancer. An estimated 170 million persons are chronically infected worldwide. The Hepatitis C virus is the pathogen agent responsible for hepatitis C. HCV is an enveloped RNA-positive virus of the *flaviviridae* family. The HCV genome shows remarkable sequence variability. This variability leads to the classification of HCV into 6 genotypes, numerous subtypes and HCV exists in each infected patient as quasi-species. The genotype may be linked to the severity of the disease and to the efficiency of the combination treatment with interferon and ribavirin. To date, no vaccine to prevent or cure HCV exists. Numerous HCV specific inhibitors have been designed and some are currently under clinical trials. However, resistances of HCV against these inhibitors have been identified. We developed the European Hepatitis C Virus Database (euHCVdb, http://euhcvdb.ibcp.fr/), a collection of functionally and structurally (3D-models) annotated HCV sequences integrated with sequence and structure analysis tools. We show below how the euHCVdb database is a useful *in silico* tool that can help drug design, combating resistance to drug treatment and understand structural biology of the HCV.

Keywords: Hepatitis C Virus, drug design, molecular modelling, database, resistance, virus, sequence analysis, structure analysis.

INTRODUCTION

Globally, 170 millions persons are chronically infected by the hepatitis C virus (HCV), a pathogen agent responsible of hepatitis C that can lead to liver cirrhosis and hepatocellular carcinoma. The HCV genome is approximately 9,600 nucleotides in length and carries a single, long open reading frame (ORF) flanked by 5' and 3' non-translated regions. The ORF encodes a polyprotein of about 3,000 amino acids that is processed by cellular and viral proteases to yield at least ten mature proteins: C, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B [1,2]. The sequence diversity among HCV genomes leads to the definition of a large number of subtypes distributed into six genotypes [3]. In addition, HCV exists within its hosts as a pool of genetically distinct but closely related variants referred to as quasispecies [4]. It is now well established that the genotype is a predictive factor of the response to interferon-alpha therapy [5]. The combination of interferon and ribavirin is the only current therapy available. Many potent HCV specific drugs have been designed and some (*e.g.* Vertex VX-950) are on late stage of clinical trials but drug resistances have been identified [6-9]. In order to help understanding and combating these drug resistance phenomena, we developed the European HCV Database (euHCVdb, http://euhcvdb.ibcp.fr/) [10]. This database combines computer-annotated HCV sequences with protein three-dimensional (3D) models and it is linked to numerous sequence and structure analysis tools on dedicated websites. The euHCVdb is mainly oriented towards the structural biology of HCV, including protein sequence, structure and function analyses in comparison with the other HCV specialized databases developed [11]. We describe in this review the data provided by euHCVdb and we illustrate how to exploit them with the integrated analysis tools provided on the database website to compare on a structural basis HCV protein variants and to analyze their interactions with drugs.

The HCV Sequence Database and Sequence Analysis Tools

The euHCVdb (current release 95.0) is updated on a monthly basis from the EMBL nucleotide sequence database [12] thanks to a fully automated annotation procedure. This procedure is an improved version of the procedure used for the French HCV database developed in 1999 [13]. The sequence annotation results in the creation of the sequence and 3D-models (euHCVdb3D) databases. The automatic annotation procedure ensures standardized nomenclature for all entries across the database and builds a description of the HCV genomic regions and proteins that are included in the entry. The sets of pre-computed multiple sequence alignments of complete proteins and reference genomes provided on the euHCVdb website are also updated monthly.

The format of a euHCVdb entry is an extension of the EMBL database format and provides: entry identifier built from the EMBL database primary accession number, bibliographic references, cross-references to external databases, genotype/subtype, well-characterized sites (*e.g.* hypervariable region 1, HVR1) or domains (*e.g.* NS3 helicase), the source of the sequence (*e.g.* isolate), and structural data as

*Address correspondence to this author at the Institut de Biologie et Chimie des Protéines; UMR5086; CNRS; Université Lyon 1; IFR128 BioSciences Lyon-Gerland; 7, passage du Vercors, 69367 Lyon CEDEX 07, France; E-mail: g.deleage@ibcp.fr
[#]*Authors have an equal contribution to works.*

1871-5265/09 $55.00+.00

3D protein models. The extension made to the EMBL database format is a set of new qualifiers that store specific data. Thus, the *genotype* (genotype deposited in EMBL entry by authors of the sequence and retrieved by text parsing from the original EMBL data), the *genotype_prov* (provisional genotype computed using a set of reference sequences, [3]) and the *genotype_conf* (confirmed genotype computed using a set of reference sequences, [3]) qualifiers are used to store the genotype data in *source* feature. For *mat_peptide* feature, we added a *prod_ft* (product feature) qualifier for protein annotation. The protein annotations are done in close collaboration with the Swiss-Prot group of the Swiss Institute of Bioinformatics. The format and the controlled vocabulary of the UniProtKB/Swiss-Prot database [14] are used in euHCVdb at the level of the *prod_ft*. A *prod_ft* is a formatted quallifier containing start/end position in the protein sequence, type of *prod_ft* (*e.g*. site, transmem, model3d) and a description text. The euHCVdb is accessible through a Web site at URL http://euhcvdb.ibcp.fr. It is divided into static and dynamic parts. In the dynamic part, a query system (not shown) allows the building of dynamic sets of sequences or 3D models according to user defined criteria (> 30 different criteria can be defined) and selected by the user through the query interface, for example: 'extract all the 3D models of the NS3 protease of provisional genotype 2'. To perform such query, the user should select *sequence type*='feature protein', *Standard name*='ns3', *Genotype/Subtype*='2*' and *Feature protein key*='model3d'. The results (Fig. **1A**) are displayed in a table where each row corresponds to one protein feature that is described by a small set of identifiers and characteristics such as accession number, protein name, genotype, isolate, length and description including template Protein DataBank (PDB, [15]) code. When a query is performed on *model3d* a link to euHCVdb3D is also mentioned in this table and allows the interactive analysis of the 3D model of the protein feature (see bleow). A toolbox is also available to export data (*e.g.* sequences) to other resources. For example, the sequences matching previous criteria can be exported to the generic integrated sequence analysis web server NPS@ [16] of the PBIL [17]. The sequences can also be extracted as a Fasta/Pearson formatted sequence database and uploaded in the HCV dedicated tools like *Numbering* [18] or *Extract*. The *Extract* tool allows the simultaneous computation of multiple sequence alignment, residues repertoire (Fig. **1B**) and Shannon entropy. The residue repertoire allows the analysis of the residue variability at each position of the alignment. This is useful to compare hundreds of variants of different genotypes/subtypes as it summarizes the alignment and highlights variable positions and the nature of the residue found while filtering for not frequently observed residue change that can be due to sequencing errors. When the query looks for 3D models, the matching entries can be exported to the 3D models database for further interactive structural analyses (Fig. **1C**).

euHCVdb3D The HCV Protein Models Database and Structure Analysis Tools

The gap between known protein sequences, deposited in the PDB [15], and solved protein structures is more and more important. Moreover, the need of structural information to better understand structure-function-interaction relationships, including protein dynamics and drug design purpose, of protein is more and more crucial. In order to reduce this gap and to increase the level of structural information, homology molecular modelling techniques have been developed. In our laboratory, we developed Modeome3D (Garnier, N. *et al.*, manuscript in preparation), an automated management system for large-scale protein 3D models generation and storage. In this process, template selection is ensured by a FastA [19] sequence similarity search or a PSI-BLAST [20] search if no satisfying template has been found by FastA. In order to validate distantly related templates, the Structural Overlap (SOV) parameter is computed [21]. We showed that the SOV parameter is able to discriminate between true and false templates [22]. Thus, templates with a SOV parameter equal or above 60 are used for modelling. The 3D models are computed with the Geno3D program for automatic comparative homology modelling of proteins [23]. All the computed models are stored in a relational database that can be queried according to several criteria and model can be visualized with an interactive interface developed around the Jmol applet [24]. Thus, the euHCVdb sequence annotation procedure ends by the euHCVdb3D model database computation with the Modeome3D system. For each sequence with a *model3d* product feature (*prod_ft* qualifier), which includes the template PDB code, a 3D model is computed is stored in the euHCVdb3D.

The 3D models of different variants (Fig. **1C**) can be analyzed using the PIG Web server for protein structure analysis (Bettler, E. *et al.*, manuscript in preparation). Selected 3D models can be superimposed to generate a multiple structural alignment and root mean square deviation (RMSD) matrix (Fig. **1D**) as well as a local RMSD. The user can interactively analyze the fitted models with the Jmol applet that is dynamically linked to the corresponding sequence structural alignment. In this way, one can identify the sequence variability between the viral variants at the 3D level using information from the *Repertoire* results (Fig. **1B**). The relevance of taking into account of sequence information for protein structure analysis is also implemented in the MAGOS server developed in our team [25]. The Web interface of MAGOS is designed for an interactive approach of structural information within the framework of the evolutionary relevance of mined and predicted sequence information. Such analyses is helpful for a better understanding of sequence-structure-function relationships which is of high relevance *e.g.* for understanding drug resistance. The user can also visualize the residue conservation at the model level, obtain a list of accessible residues, or detect ligand-binding or active sites using the SuMo Web server [26]. All these tools have been used to further investigate the VX-950 interaction with HCV NS3-4A protease resistance and impact of mutation.

In Silico Analysis of NS3-4A Mutants for VX_950 Interaction

The VX-950 molecule for vertex (R1) is a specific inhibitor of the NS3-4A protease of the HCV virus [27]. The VX-950 molecule is a peptide modified at the level of backbone and side chain. It occupies the catalytic site of the

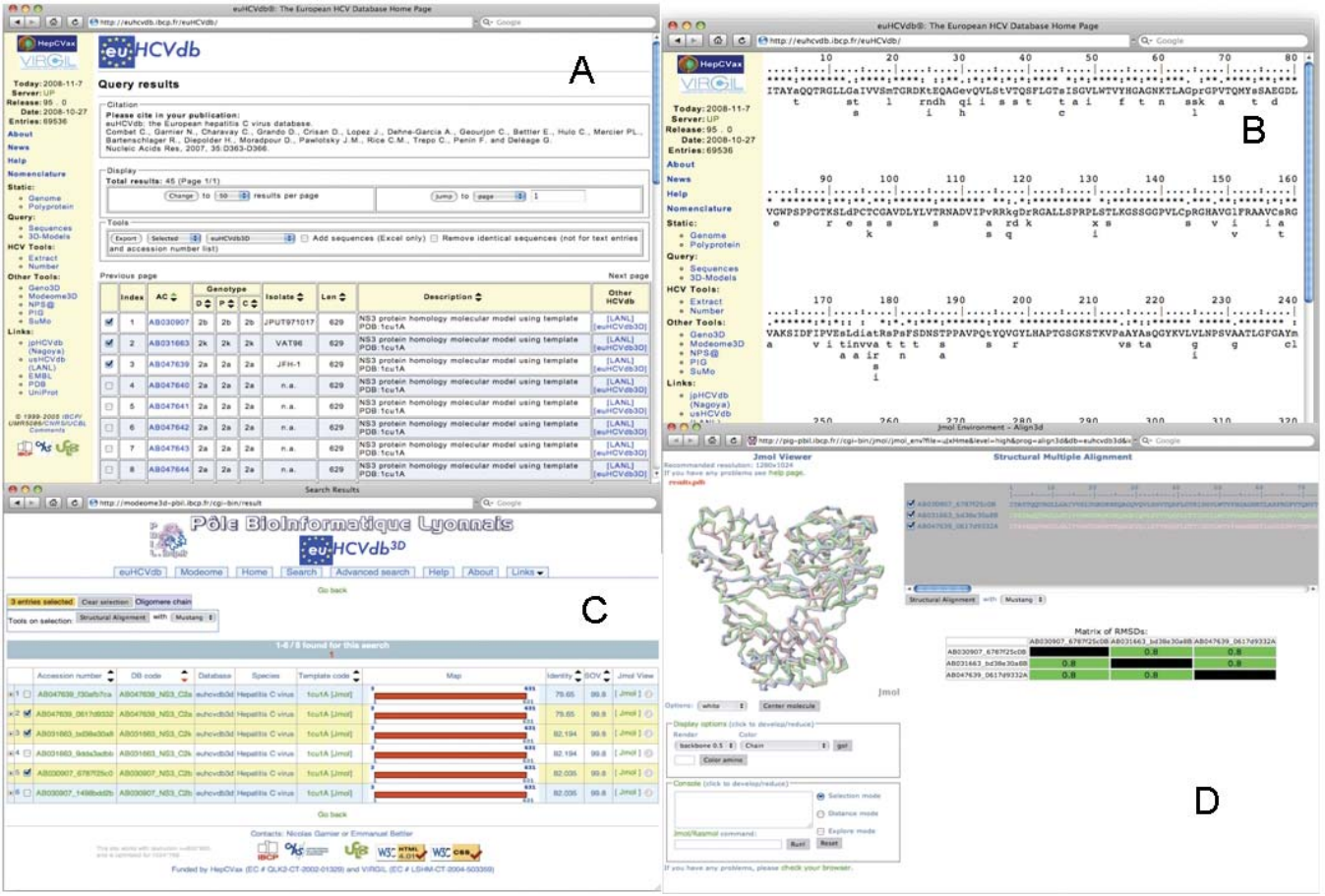

Fig. (1). Example of analysis with euHCVdb. A) Result of a request searching for NS3 protein of genotype 2 and for which a 3D model exists. B) Repertoire of amino-acid residues computed with all the NS3 sequences matching the previous criteria. C) Result of export of the three selected sequences to the 3D-models database D) Interactive structural analysis of the the three selected NS3 models showing molecule (left), multiple sequence alignment (top right) and overall RMSD value between models (right bottom).

NS3-4A protease and is active at the nanomolar scale. The binding to the protease molecule was investigated by NMR and the binding site of the NS3-4A protease was divided into 4 different pockets termed S1 to S4. These pockets are mostly hydrophobic. The amino acid side chains were optimized to increase the selectivity and the binding energy of the VX-950 molecule. The pocket S4, exhibits a solvent access surface mostly represented by (or dependent from) V158. This pocket is the binding site of the cyclohexane moiety of the VX-950 molecule. The surface of the S3 pocket is determined by (or dependent from) the V132 and the VX-950 molecule brings the t-Butyl moiety on contact. The depth of the backbone on the structure of the protease forms the S2 pocket and additional cycle was added to the proline residues to increase the hydrophobic properties of the moiety. The S1 pocket is composed by the other side of the V132 and the hydrophobic part of the side chain of the K136. The VX-950 molecule was optimised with a propyl moiety for this pocket. All the side chains were optimized to increase the binding energy of the molecule and the different residues formed hydrogen bonds with the pseudo backbone of the protease molecule.

In order to evaluate the impact of a single mutation on the efficiency of the drug an approach using the 3D model of the targeted protein was developed.

Starting from the euHCVdb3D database, the 3D model of the NS3-4A protease with the entry AF009606 from genotype 1a was downloaded. Sequences of the model of the complex NS3-4A start from I32 to S181 for NS3 and G21 to S32 for NS4A. Three different mutants at position A156 were created by using Sybyl molecular modelling package [28], the A156S, A156T and A156V. Mutants at position A156 are known to exhibit resistance to inhibition by the VX-950 molecule. The native model and the three mutant models were prepared for molecular dynamics simulations. These simulations were carried out by using Amber 9 with the Amber force field 03. The simulations were made in a solvent box filled with TIP3P water molecules with a minimum distance from the limit of the periodic boundary condition to the model of 12 Å. A Zinc ion was added to each model in the Zinc fingers domain and global charge was neutralized using chloride ions. Parameters for the different ions were obtained from Jorgensen *et al.* [29]. The simulation used an integration step of 2fs with the shake algorithms and the Particle Mesh Ewald to compute

electrostatic contributions. All simulations were carried out for 5 ns simulation time with a conformational sample every 20 ps. The system was heated from 100K to 300K in 40 ps and reached equilibrium at 500 ps.

Analyses of all four dynamics were performed with the VMD software [30]. Three different analyses were performed. A 2D Root Mean Square Deviation (RMSD) on every heavy atoms of the structure was computed. For each couple the RMS was placed on a 2 graph XY plot in false colours. The colour code is white from 0 Å to black for RMSD greater than 10 Å. The analysis of the four 2D-RMSD (Fig. **2**) revealed that the AF009606-A156S had a different 2D-RMS diagram in comparison to the three others (Native (AF009606), A156V and A156T). For the three last models no real change was observed on the structure and the RMSD remained at low level. The three models exhibited a slow and weak evolution and every conformational cluster was strongly correlated to the previous one. The analysis of the 2D-RMSD of the AF009606-A156S is quite different: The evolution is drastic, and there are no correlation between the starting point and the main conformational cluster (frame roughly 25 to 75). The structure makes its conformational evolution via a "transitional" first cluster and then makes a quick evolution which is conserved up to the end of the simulation. The last 10 frames show that, the structure makes a small evolution which looks like different from the starting point of the main cluster. In conclusion of these analyses, the first three models do not really evolve or undergo a real change while the AF009606-A156S makes a drastic evolution and is still evolving at the end of the simulation. We can deduce that substituting the alanine residue at position 156 by a serine has important consequences on the structure.

In order to complete a more detailed analysis we made a script to measure the local RMSD on the heavy atoms of five consecutive residues for all sampled frames and affected the

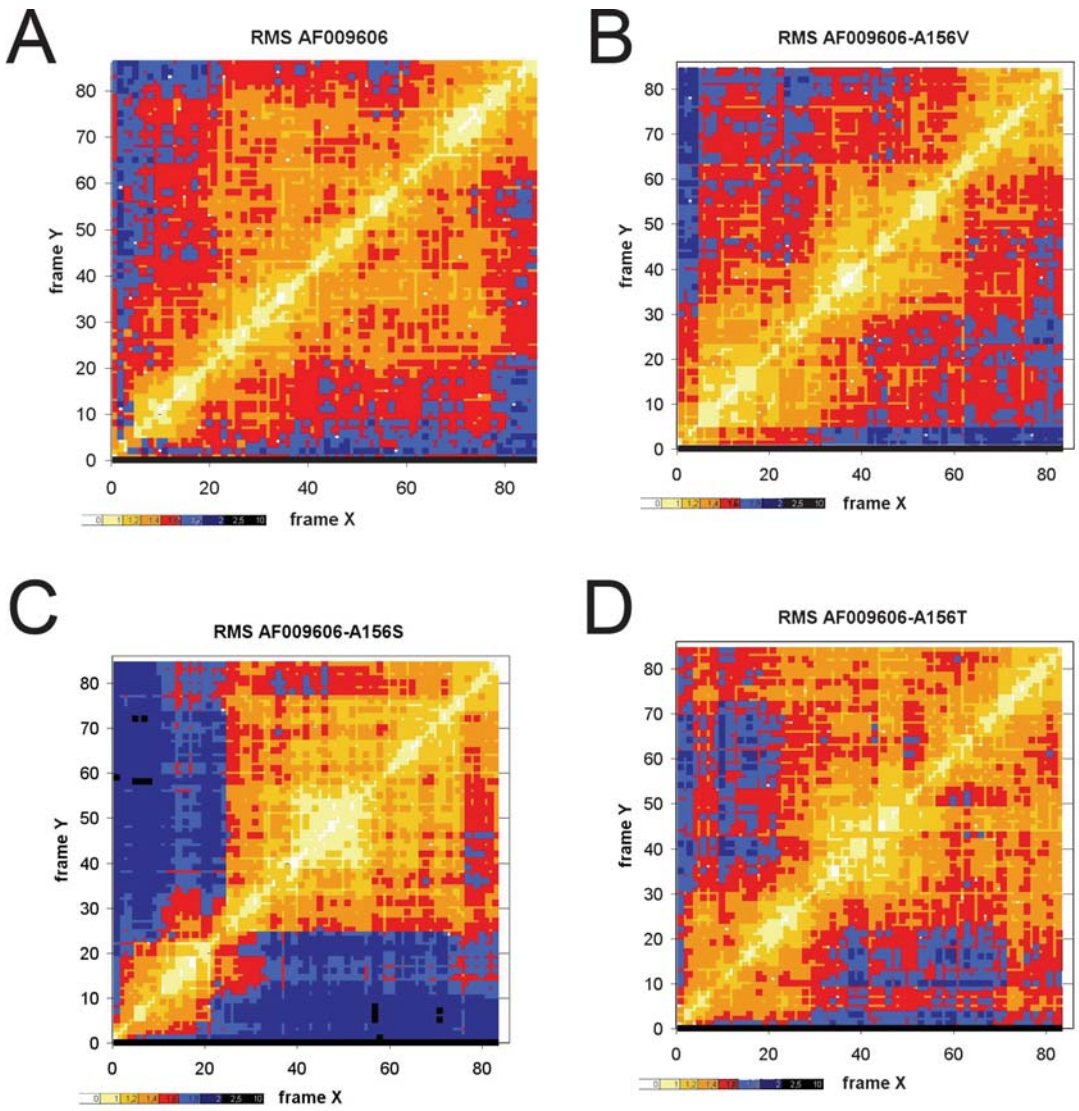

Fig. (2). 2D-RMSD of the trajectories of the native structure and the 3 mutants. A frame is sampled each 100 ps for 2D-RMSD. Frames are not sampled during the first ns of the simulation because the system doesn't reach its equilibrium. False colours were applied in function of RMSD for each couple of frame.

mean RMSD values to the central residue. All frames were aligned on the 5 amino acids involved in the calculation. These calculations were performed for all residues by sliding the centred window to the next residue. The RMSD window by residues of the dynamics of the native structure was defined as the standard and the value was deduced from the values of all mutants for the same residues. The final values for mutants were taken as absolute values. The graphs (Fig. **3**) clearly show which area exhibits an abnormal value in comparison to the native structure (RMSD = 0).

A variation value of 3-4 Å is not relevant as deduced from the comparison of two stochastic simulations. The analysis of the Fig. (**3**) reveals that there are very important variations of AF0096006-A156S on residues 165-172 and in particular 8 Å for residue 170. There is also a significant variation of about 4.3 Å on residue 101 of AF009606-A156V. These two significant variations on RMSD must be analysed to provide information on the potential resistance mechanism caused by a single mutation and the affected "suburb" of the catalytic site.

If the two last analyses could be fully automated and were able to become a part of a set of analytical tools, then the last analyses would provide a classical analysis of the area to understand modifications. The analysis of the trajectory of AF0096006-A156S and the 3D structure of the ligand binding pocket filled with VX-950 (Fig. **4b**) reveal that the S156 forms a hydrogen bond with S139. The S139 is localized on a loop adjacent to the S156 which is in the middle of one section of a β sheet. In the native structure (Fig. **4a**) the distance between the Carbon beta of the A156 and the oxygen of the S139 is about 7 Å versus 3 in the A156S. This section is engaged in a β barrel scheme and is connected to the next sheet by a loop which is a part of the catalytic site. The distance between the two residues becomes shorter than that of the native structure which has no hydrogen bond. This small distance contributes to destabilize the other sheet and also consequently the connection loop. This mechanism could explain the resistance to the VX-950 molecule of the mutant A156S. The other modification centred on residue 101 of the AF0096006-A156V is less important because the value is only 4.3 Å. The S101 is located in the middle of a solvent accessible loop which originates from the first sheet of the previously discussed β barrel.

Molecular dynamics simulations of mutants compared to their native structure could offer a new tool to evaluate the impact of the mutations. This tool could be fully automated and added to the euHCVdb3D computation and analysis interface. However, the calculation for one single mutation is time consuming. It takes roughly one week on 2 dual core CPU. A possible strategy could be to use the remote access to large cluster computer or to distributed systems such as grid computer.

CONCLUSION

The euHCVdb database provides useful data that are a starting point for further analysis with integrated tools provided on the database website or with local tools in the case of drug design or molecular dynamics. The euHCVdb database has been running since March 2005. The current release number 95 (October 2008) comprises 69,536 entries and 11,209 protein 3D models representing 6,650 protein chains for a total of 3,429,197 residues. The database

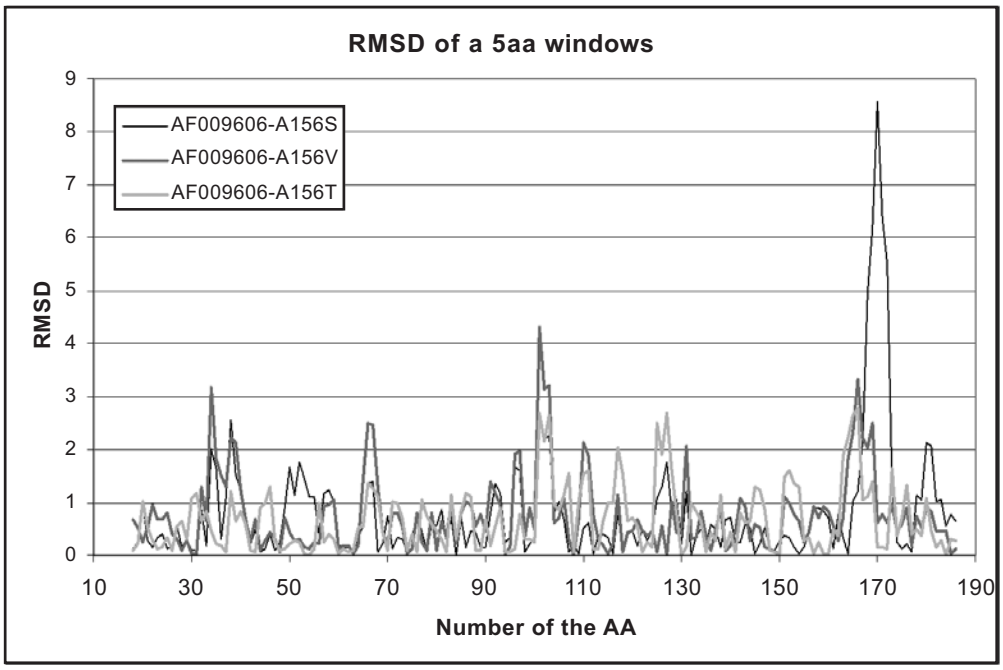

Fig. (3). Variation of the RMSD by residue compared from the native structure to different mutants. Values reported to the central residue of a window of 5 residues long along the molecular dynamic simulation.

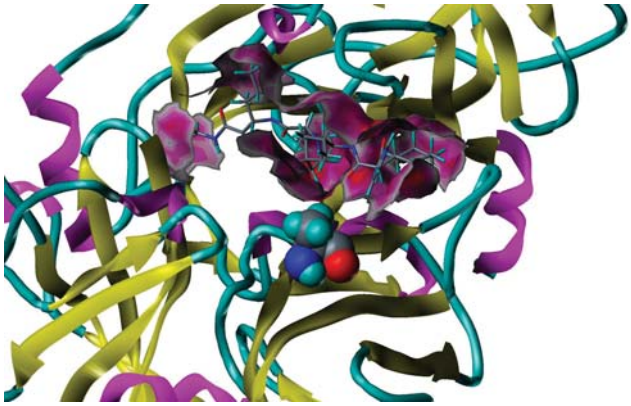

Fig. (4a). VX 950 in interaction with the model of the NS3/4A sequence AF009606. The solvent access surface of the binding site on the NS3-4A is displayed with distance function colour, close distance: red colour / longer distance: gray. The VX950 molecule is displayed in stick mode, the A156 is displayed with spacefill atoms. The rest of protein is displayed with a ribbon.

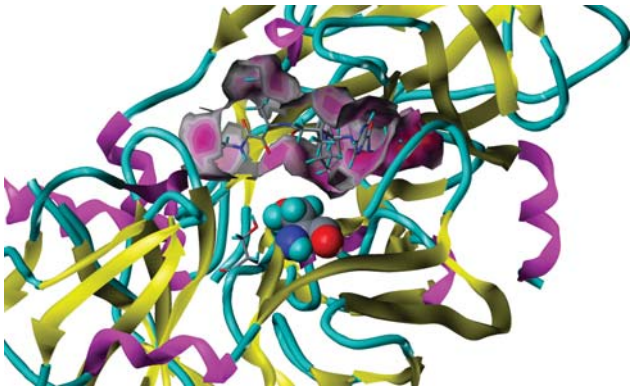

Fig. (4b). VX 950 in interaction with the model of the NS3/4A sequence AF009606-A156S. The solvent access surface of the binding site on the NS3-4A is displayed with distance function colour, close distance: red colour / longer distance: gray. The VX950 molecule is displayed in stick mode, the S156 is displayed with spacefill atoms and the S139 in stick mode. The rest of protein is displayed with a ribbon.

currently receives about 5000 queries per month. A 3D model database euHCVdb3D has also been developed for the generation, the storage and the management of 3D protein models. These models can be used for molecular dynamics simulations to compare mutants with wild structure. The euHCVdb suite of programs offers *in silico* tools to evaluate the structural impact of mutations. In this paper we have illustatred the use of these tools to investigate the structural response to the binding capability of VX-950 molecule to mutants. The evaluation of the selectivity and the resistance to a molecule, is becoming the cutting edge in molecular modelling / bioinformatics sciences.

ACKNOWLEDGMENTS

The European HCV Database is funded by the European Commission (HepCVax Cluster, FP5 grant QLK2-2002-01329 and VIRGIL Network of Excellence, FP6 grant LSHM-CT-2004-503359) and Agence Nationale de Recherches sur le SIDA et les Hépatites Virales (ANRS). Nicolas Garnier is the recipient of an Association Française contre les Myopathies (AFM) doctoral fellowship. Funding to pay the Open Access publication charges for this article was provided by European Commission (VIRGIL Network of Excellence, grant LSHM-CT-2004-503359).

REFERENCES

[1] Lindenbach, B.D.; Rice, C.M. *Nature*, **2005**, *436*, 933-938.
[2] Penin, F.; Dubuisson, J.; Rey, F.A.; Moradpour, D.; Pawlotsky, J.M. *Hepatology*, **2004**, *39*, 5-19.
[3] Simmonds, P.; Bukh, J.; Combet, C.; Deléage, G.; Enomoto, N.; Feinstone, S.; Halfon, P.; Inchauspe, G.; Kuiken, C.; Maertens, G.; Mizokami, M.; Murphy D.G.; Okamoto, H.; Pawlotsky, J.M.; Penin, F.; Sablon, E.; Shin-I, T.; Stuyver, L.J., Thiel, H.J.; Viazov, S.; Weiner A.J.; Widell, A. *Hepatology*, **2005**, 42, 962-973.
[4] Weiner, A.J.; Christopherson, C.; Hall, J.E.; Bonino, F.; Saracco, G.; Brunetto, M.R.; Crawford, K.; Marion, C.D.; Crawford, K.A.; Venkatakrishna, S; Miyamura, T.; McHutchinson J., Cuypers T., Houghton M. *J. Hepatol*., **1991**, *13* (Suppl 4), S6-S14.
[5] Pawlotsky, J.M. *Clin. Liver Dis*., **2003**, *7*, 45-66.
[6] Lin, C.; Kwong, A.D.; Perni, R.B. *Infect. Disord. Drug Targets*, **2006**, *6*, 3-16.
[7] Koev, G.; Kati W. *Expert Opin. Investig. Drugs*, **2008**, *17*, 303-319.
[8] Liu-Young, G.; Kozal, M.J. *AIDS Patient Care STDS*, **2008**, *22*, 449-457.
[9] Kim, A.Y.; Timm *J. Expert Rev. Anti. Infect. Ther*., **2008**, *6*, 463-478.
[10] Combet, C.; Garnier, N.; Charavay, C.; Grando, D.; Crisan, D.; Lopez, J.; Dehne-Garcia, A.; Geourjon, C.; Bettler, E.; Hulo, C.; Le Mercier, P.; Bartenschlager, R.; Diepolder, H.; Moradpour, D.; Pawlotsky, J.M.; Rice, C.M.; Trépo, C.; Penin, F.; Deléage, G. *Nucleic Acids Res*., **2007**, *35*, D363-D366.
[11] Kuiken, C.; Mizokami, M.; Deléage, G.; Yusim, K.; Penin, F.; Shin, I.T.; Charavay, C.; Tao, N.; Crisan, D.; Grando, D.; Dalwani, A.; Geourjon, C.; Agrawal A.; Combet, C. *Hepatology*, **2006**, *43*, 1157-1165.
[12] Cochrane, G.; Akhtar, R.; Aldebert, P.; Althorpe, N.; Baldwin, A.; Bates, K.; Bhattacharyya, S.; Bonfield, J.; Bower, L.; Browne, P.; Castro, M.; Cox, T.; Demiralp, F.; Eberhardt, R.; Faruque, N.; Hoad, G.; Jang, M.; Kulikova, T.; Labarga, A.; Leinonen, R.; Leonard, S.; Lin, Q.; Lopez, R.; Lorenc, D.; McWilliam, H.; Mukherjee, G.; Nardone, F.; Plaister, S.; Robinson, S.; Sobhany, S.; Vaughan, R.; Wu, D.; Zhu, W.; Apweiler, R.; Hubbard T.; Birney, E. *Nucleic Acids Res.*, **2008**, *36*, D5-D12.
[13] Combet, C.; Penin, F.; Geourjon, C.; Deléage, G. *Appl. Bioinformatics*, **2004**, *3*, 237-240.
[14] The UniProt Consortium. *Nucleic Acids Res.*, **2006**, *34*, D187-191.
[15] Deshpande, N., Addess, K.J., Bluhm, W.F., Merino-Ott, J.C., Townsend-Merino, W., Zhang, Q., Knezevich, C., Xie, L., Chen, L., Feng, Z.; Green, R.K.; Flippen-Anderson, J.L.; Westbrook, J.; Berman, H.M.; Bourne, P.E. *Nucleic Acids Res.*, **2005**, *33*, D233-D237.
[16] Combet, C.; Blanchet, C.; Geourjon, C.; Deléage, G. *Trends Biochem. Sci*., **2000**, *25*, 147-150.
[17] Perrière, G.; Combet, C.; Penel, S.; Blanchet, C.; Thioulouse, J.; Geourjon, C.; Grassot, J.; Charavay, C.; Gouy, M.; Duret, L.; Deléage, G. *Nucleic Acids Res*., **2003**, *31*, 3393-3399.
[18] Kuiken, C.; Combet, C.; Bukh, J.; Shin-I, T.; Deléage, G.; Mizokami, M.; Richardson, R.; Sablon, E.; Yusim, K.; Pawlotsky, J.M.; Simmonds, P.; Los Alamos HIV Database Group. *Hepatology*, **2006**, *44*, 1355-1361.
[19] Pearson, W.R.; Lipman, D. J. *Proc. Natl. Acad. Sci. USA*, **1988**, *85*, 2444-2448.

[20] Stojmirović, A., Gertz, E.M.; Altschul, S.F.; Yu, Y.K. *Bioinformatics*, **2008**, *24*, i15-23.
[21] Zemla, A.; Venclovas, C.; Fidelis, K.; Rost, B. *Proteins*, **1999**, *34*, 220-223.
[22] Geourjon, C.; Combet, C.; Blanchet, C.; Deléage, G. *Protein Sci.*, **2001**, *10*, 788-797.
[23] Combet, C.; Jambon, M.; Deléage, G.; Geourjon, C. *Bioinformatics*, **2002**, *18*, 213-214.
[24] Jmol: http://wiki.jmol.org
[25] Garnier, N.; Friedrich, A.; Bolze, R.; Bettler, E.; Moulinier, L.; Geourjon, C.; Thompson, J.D.; Deléage, G.; Poch, O. *Bioinformatics*, **2006**, *22*, 2164-2165.
[26] Jambon, M.; Andrieu, O.; Combet, C.; Deléage, G.; Delfaud, F.; Geourjon, C.; *Bioinformatics*, **2005**, *21*, 3929-3930.
[27] Lin, K.; Perni, R.B.; Kwong, A.D.; Lin, C. *Antimicrob. Agents Chemother.*, **2006**, *50*, 1813-1822,
[28] Case, D.A.; Darden, T.A.; Cheatham, T.E.; Simmerling, C.L.; Wang, J.; Duke, R.E.; Luo, R.; Merz, K.M.; Pearlman, D.A.; Crowley, M.; Walker, R.C.; Zhang,W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra,G.; Wong, K.F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D.H.; Schafmeister, C.; Ross, W.S.; Kollman, P.A. AMBER 9, University of California, San Francisco **2006**.
[29] Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Imprey, R.W.; Klein, M.L. *J. Chem. Phys.*, **1983**, *79*, 926.
[30] Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.*, **1996**, *14.1*, 33-38.

Received: November 21, 2008 Revised: February 5, 2009 Accepted: February 5, 2009