



HAL
open science

Multisensory VR interaction for protein-docking in the CoRSAIRE project

Nicolas Férey, Julien Nelson, Christine Martin, Lorenzo Picinali, Guillaume Bouyer, Alex Tek, Patrick Bourdot, Jean Marie Burkhardt, Brian F,g Katz, Mehdi Ammi, et al.

► **To cite this version:**

Nicolas Férey, Julien Nelson, Christine Martin, Lorenzo Picinali, Guillaume Bouyer, et al.. Multisensory VR interaction for protein-docking in the CoRSAIRE project. *Virtual Reality*, 2009, 13 (4), pp.273-293. 10.1007/s10055-009-0136-z . hal-01753897

HAL Id: hal-01753897

<https://hal.science/hal-01753897v1>

Submitted on 30 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Multisensory VR interaction for Protein-Docking in the CoRSAIRe project

N. Férey · J. Nelson · C. Martin · L.
Picinali · G. Bouyer · A. Tek · P. Bourdot ·
J.M. Burkhardt · B.F.G Katz · M. Ammi ·
C. Etchebest · L. Autin

Received: date / Accepted: date

Abstract Proteins assume their function in the cell by interacting with other proteins or biomolecular complexes. To study this process, computational methods, called protein docking, is used to predict the position and orientation of a protein ligand when it is bound to a protein receptor or enzyme, taking into account chemical or physical criteria. This process is intensively studied in order to discover new protein biological functions and to better understand how these macromolecules assume these functions at the molecular scale. Pharmaceutical research also employs docking techniques for a variety of purposes, most notably in the virtual screening of large databases of available chemicals in order to select likely drug candidates. The basic hypothesis of our work is that Virtual Reality and multimodal interaction can increase efficiency in reaching and analysing docking solutions, complementarily to fully computational docking approach. To this end, we conducted an ergonomic analysis of the protein-protein current docking task. Using these results, we designed an immersive and multimodal application where Virtual Reality devices, such as 3D mouse and haptic device, are used to interactively manipulate two proteins for exploring possible docking solutions. During this exploration, visual, audio and haptic feedbacks are combined to render and evaluate chemical or physical properties of the current docking configuration.

N. Férey, C. Martin, G. Bouyer, A. Tek, P. Bourdot, B.F.G Katz, M. Ammi
Laboratoire d'Informatique et de Mécanique pour les Sciences de l'Ingénieur, Université Paris XI BP133, 91403 Orsay Cedex, France
E-mail: tek,pb,ammi,katz@limsi.fr, nicolas.ferey@ibpc.fr, guillaume.bouyer@ensiie.fr

J. Nelson, J.M. Burkhardt
Laboratoire Ergonomie-Comportement-Interactions, Université Paris V, 45, rue des Saints-Peres, 75006 Paris, France
E-mail: jean-marie.burkhardt, julien.nelson@univ-paris5.fr

L. Autin, C. Etchebest
Institut National de la Santé et de la Recherche Médicale, Université Paris VII, INTS, 6, rue Alexandre Cabanel, 75739 Paris Cedex 15, France
E-mail: catherine.etchebest@univ-paris-diderot.fr, ludovic.autin@gmail.com

L. Picinali
Institut de Recherche Acoustique et Musical, 1, place Igor Stravinsky, 75004 Paris, France
E-mail: lorenzo@limsi.fr

Keywords Protein Docking · User-Centered Design · Virtual Reality · Multimodal Rendering

1 Introduction

Protein-protein docking refers to the structural biology problem consisting of predicting how proteins bind in order to make up functional complexes within the cell, based on the 3D structure of proteins and on their physicochemical properties (see section 2). Knowledge of the structure of protein-protein complexes allows scientists to better understand the key mechanisms at work in protein-protein interaction. This is a major scientific bottleneck, in terms of both theoretical (understanding a protein function) and applied research (specific inhibition of a protein function for drug design).

Current methods for protein-protein docking include (1) automatic stages (explained in section 3.1) which take into account proteins topology, as well as energy (*i.e.* physicochemical) properties and (2) stages of molecular visualization allowing evaluation of the results. The automatic stage is costly in terms of processing time, and yields a large number of docking configurations which can't be discriminated using objective and automatic parameters. Therefore, a manual stage of analysis is also necessary, requiring visualization of the results by an expert. However, this visual analysis requires large amounts of information to be processed simultaneously by the docking expert: manipulation of 3D objects, physicochemical data, biological data, etc.

Given the current limitations of docking tools, it seems relevant to develop complementary or alternative approaches to docking. In project CoRSAIRe (Combination of Sensorimotor Renderings for the Immersive Analysis of Results) our hypothesis is that using Virtual Reality (VR) technologies and related multi-sensorimotor interactions may help experts in the docking task. There are several reasons for this. Firstly, stereoscopy, especially when it is adaptative, may improve perception of 3D protein models. Furthermore, direct manipulation of several proteins at the same time, afforded by peripherals commonly used today for such tasks (*e.g.* 3D mice, force-feedback interfaces, etc.) may be more intuitive and efficient than traditional, desktop WIMP¹ interface. Additionally, multimodal management of sensorimotor feedbacks (using an approach aiming to dynamically specify adaptation of visual, haptic and audio renderings to characteristics of information at use) is one possible answer to the problems linked to simultaneous presentation of large amounts of data. Finally, a strongly interactive approach of VR docking allows the docking expert to be placed on the forefront of the work, rather than allowing an automatic algorithm complete control over the generation of possible sets of solutions. We believe our approach, which combines benefits of multimodal interaction, will allow capitalization of docking experts occupational skills (in biology, crystallography, bioinformatics) modelling in order to improve speed of predictions regarding the structure of protein-protein complexes, as well as search efficiency and quality of results in analyzing possible solutions.

Applications for interactive docking, whether multimodal or immersive, have already developed in the past and are presented in section 3.2. However, this work seems set back by a lack of clear knowledge regarding user needs and working practices. In order to gather this, we conducted ergonomic analyses of the protein docking task as carried out by experts in the field today, in order to determine advantages and

¹ Acronym of Window, Icon, Menu, and Pointing device.

drawbacks of existing tools and help design an innovative and relevant paradigm for multimodal VR interaction. The data, recommendations and task scenarii based on these analyses (see section 4) allowed us to model a complete hybrid approach to docking, combining interactive and automatic stages. This approach was then implemented within a first prototype. We then present its architecture, related constraints and the technical solutions chosen to circumvent them (in section 5.1). We are currently in the course of evaluating this prototype in collaboration with docking experts (see section 6).

2 Protein-protein docking context

Proteins can be viewed both as the building blocks and workforce of cells. They are synthesized based on a portion of DNA (Deoxyribo-Nucleic Acid) called a coding sequence or gene. It is then transcribed in the form of mRNA (Messenger RiboNucleic Acid). This mRNA is then translated by ribosomes in the form of a protein, based on a specific coding scheme (figure 1A). Each triplet of mRNA bases corresponds to one AA (Amino Acid) or residue. There are twenty basic types of AAs which differ according to the chemical nature of its side chain, named R. It is therefore possible to classify AAs in groups which share specific characteristics, which are inductive of various types of behavior, *e.g.* non-polar (hydrophobic) vs. polar (hydrophilic) groups, which are respectively electrically neutral and electrically charged. These various physicochemical properties give rise to interactions at the atomic level, inducing protein folding which contributes to its stability (figure 1B). These properties also play a crucial part in protein-protein interactions.

Proteins, therefore, are long chains composed of successive amino acids folded in space, which are the product of the expression of an organism's genetic makeup. But in order to execute their functions within cells, proteins must undergo folding and take a specific 3D form. This form may be characterized according to four structural levels (see figure 1B). The order in which residues are linearly arranged, *i.e.* their sequence, constitutes the protein's primary structure. (see figure 1B-a). Some of the structure's segments organize themselves into sequences of specific substructures called secondary structures (see figure 1B-b). These structures, stabilized by hydrogen bonds, can be divided into two groups: regular secondary structures, called alpha helices and beta sheets, which are linked together by irregular structures called loops. The arrangement of these secondary structures thus constitutes the 3D, or tertiary structure (see figure 1B-c) which determines protein function within the cell.

Once folded, proteins carry out diverse functions within the cell, such as transporting molecules to and from various components of the organism (*e.g.* hemoglobin, chaperone proteins), inter- and intracellular signaling and communications (*e.g.* hormones, neurotransmitters, ions), immune defense functions (immunoglobulins, adhesion molecules) or cellular metabolism (chlorophyll, apoptosis proteins, transcription factors, ATP synthesis). These cellular functions are closely linked to the protein's tertiary structure, but also to its interactions with other proteins. These interactions produce new entities called protein or supramolecular complexes (see figure 2). Such assemblies make up the protein's quaternary structure (figure 1B-d).

In sum, a better understanding of protein-protein interactions is a major stake for biomedical research. Indeed, designing new drugs increasingly involves targeting specific protein-protein interactions [Villoutreix et al., 2008], or alternately, involve syn-

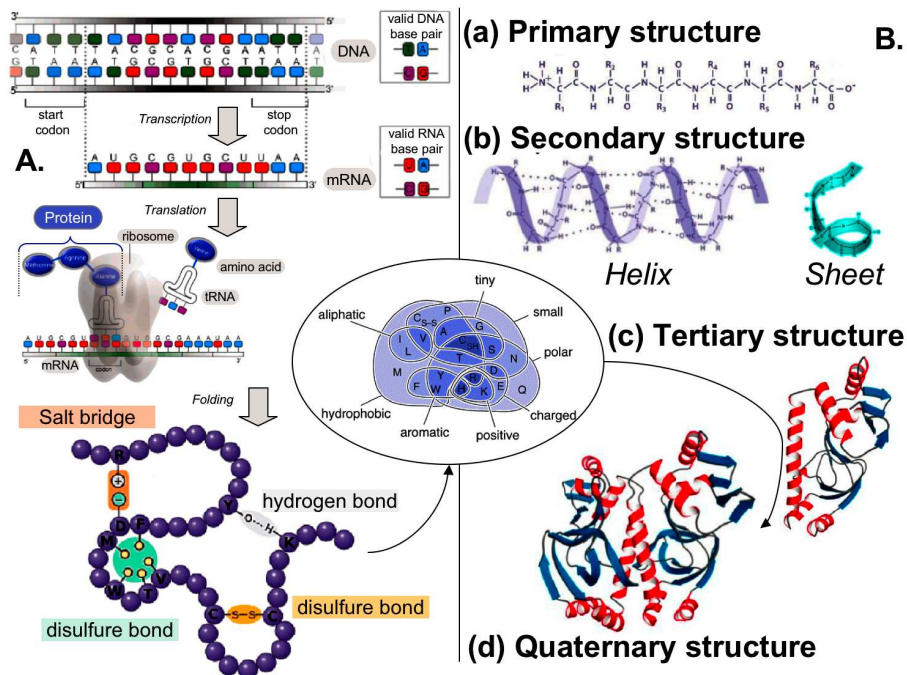


Fig. 1 (A) - Overall view of protein synthesis: transcription of DNA to messenger RNA (mRNA) and translation of mRNA to sequences of amino acids chosen from 20 possible varieties, here shown according to their physicochemical properties (using a Venn diagram). **(B)** - Based on the chemical nature of component amino acids, resulting interactions cause the protein to fold up in space. This 3D shape can be described according to four levels: (a) primary, (b) secondary, (c) tertiary and lastly (d) quaternary.

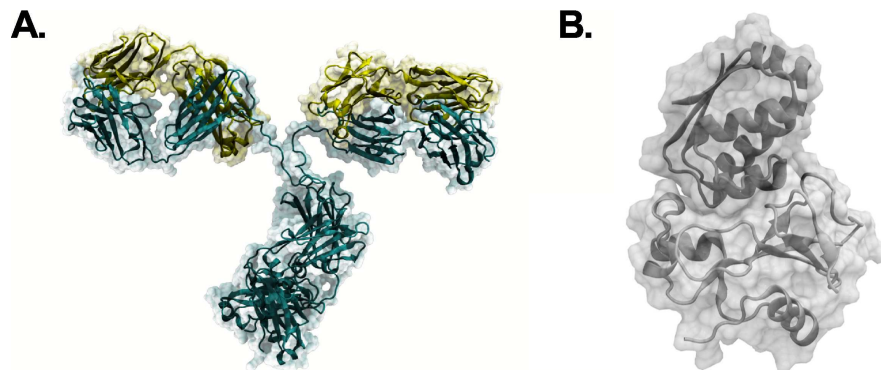


Fig. 2 (A) - Quaternary structure of an antibody (immunoglobulin), comprising a heavy chain and a light chain. This assembly allows antibodies to recognize antigens in foreign bodies in order to form a supramolecular complex. **(B)** - This complex is formed by interactions between an enzyme, barnase (below) and its substrate, barstar (above). Molecular structures are represented along with a transparent molecular surface in order to illustrate surface complementarity.

thesizing recombinant proteins meant to emulate interaction with the original native protein [Pipe, 2008]. It grows more and more necessary, therefore, to identify the 3D structure of protein complexes. Two experimental methods currently exist, allowing determination of the 3D structure of a protein complex. These are X-ray crystallography and Nuclear Magnetic Resonance (NMR). All known protein structures are currently housed on the website of the Protein Data Bank (PDB) [Berman et al., 2000]. This database contains about 50 000 protein structures for all kinds of organisms. However, this number remains small in comparison to estimates of the number of existing proteins in the natural world (*e.g.* about 20 000 proteins for Man). This is because experimental determination of protein structure is often difficult, and in some cases impossible. Indeed, solving a problem of this kind involves mass production of the protein, its purification, and in the case of crystallography, production of diffractive crystals. Furthermore, it is impossible to obtain usable structures with an acceptable resolution in the case of transmembrane proteins. In determining the complex structures, difficulties in production and purification are all the more critical, since proteins must be produced simultaneously for complexes to form. Additionally, the time necessary for crystallization may be incompatible with the lifespan of some complexes. For those reasons, numerous scientists have attempted to predict the structure of such complexes using computing tools through methods and algorithms for molecular docking.

3 Related works

In this section we present a state of the art in computer-based approaches and existing VR solutions for protein docking, then we introduce the main focuses of the CoRSAIRe project in multimodal VR interaction in that field of application.

3.1 Automatic approaches for docking

Current techniques for experimental study of the 3D structure of complexes (crystallography, NMR, electronic cryomicroscopy, etc.) have several limitations (in terms of size and type of proteins) and are costly both in terms of time and money. For that reason, computer-based (*in silico*) docking methods were developed, aiming to deduce the functional 3D structure of a complex based on single molecules, which turns out to be considerably easier and cheaper to carry out than experimental, *in vitro* methods [Grosdidier, 2007]. Current approaches are strictly computational and results are evaluated using visualization tools. These approaches can be divided into four to five successive stages (figure 3): (1) choice of the mode of representation of proteins (atomic view, pseudo-atoms, grid, etc.); (2) conformational exploration (hard docking: position and orientation of ligand; soft docking: position, orientation and shape of ligand); (3) minimization of the function to evaluate binding energy (*score*) for conformations derived from the exploration; (4) grouping by similarity and classification through evaluation or fine-tuning of the *scores*, along with a manual stage of visualization when score alone doesn't allow the native conformation to be discriminated against other generated conformations; (5) an optional stage of fine-tuning of selected complexes, through energy minimization or molecular dynamics.

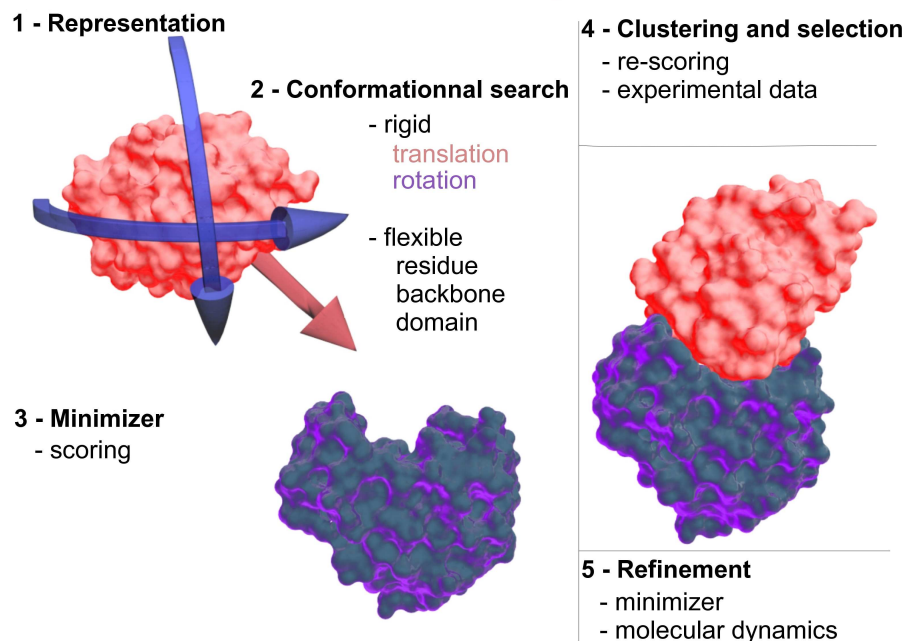


Fig. 3 The 5 stages of the docking task

A large number of docking algorithms depend on an exhaustive approach to conformational exploration, the main problem being combinatorial exploration of the number of possible solutions. These approaches can be sorted into three categories: approaches based on systematic sampling, molecular dynamics techniques and classification of modes of interaction between proteins. An ideal objective function would yield, for a given mode of interaction, the energy binding the two proteins in a complex (see section 5.2.2). Such functions aim to reproduce experimental values of free binding energy, and through minimization, reach the global minimum for all conformations of protein-protein complexes.

Consequently, in real life cases, automatic docking processes must manage two difficulties in order to reach a relevant result. The first is to process a space of potential solutions which increases in size along with the number of degrees of freedom in describing protein position and conformation, and cannot be processed in an acceptable amount of time. The second problem is that search algorithms produce local minima, and cannot easily find the global minima that is associated to the native form of the complex [Wang et al., 2003].

In order to finalize a docking simulation, experts rely upon a manual stage of visualization to analyse the generated complexes. This activity consists in a detailed analysis of residues and atoms involved in each complex interface, through the observation of hydrogen bonds, salt bridges, and especially the presence of hotspots, *i.e.* amino acids located in the interface, known to be part of the interface, according to experimental studies. However, it can prove difficult to manipulate two 3D structures simultaneously to observe the interface using traditional interaction tools, since one protein usually hides the other. That is why we believe docking assisted by VR and

multimodal interaction presents a relevant alternative to improve the work of experts in the field. Such techniques might allow a more intuitive interaction with 3D protein structures.

Finally, regarding “thinning the herd” of selected complexes, two approaches are used. One consists in minimizing the rigid bodies and lateral chains of amino acids present at the interface. This approach is implemented in several applications such as *ICM-DISCO* [Fernandez-Recio et al., 2003], *FireDock* [Andrusier et al., 2007], *PELE* [Borrelli et al., 2005], *MMTK* [Hinsen, 2000], *ATTRACT* [Zacharias, 2005], etc. The other approach involves studying the dynamic behavior of the selected complex. The software program *Gromacs* [Hess et al., 2008], for example, allows evaluation of atomic positions in time based on their physicochemical properties. This approach allows first to evaluate complex stability, and second to evaluate possible conformational changes induced by the interaction, *e.g.* loop deformation. We should add, however, that this approach remains very costly in terms of processing time, in relation to minimizers which allow to process a given configuration very quickly.

3.2 Interactive and immersive approaches for docking

Given the drawbacks of classical docking algorithms and of the new possibilities afforded by VR, several teams have taken an interest in recent years regarding problems linked to interactive and immersive docking. Early work in the field primarily involved identification of technical needs and limitations to achieve this. The *STALK* system [Levine et al., 1997] uses parallel and distributed processing, firstly to process visual renderings of the 3D protein models, and secondly to execute the algorithms for generating possible solutions. Using this system, the docking expert may visualize two proteins, assist the docking algorithm, suspend it to move a molecule using a 3D mouse, and resume the algorithm using this new position. System evaluations were concerned with comparing binding energies of solutions obtained. Another system, *VRDD* [Anderson & Weng, 1999], implemented a search algorithm using a Monte Carlo method, which allowed gradual reduction of the search space through the exploration of similar solutions. Evaluations were carried out over three test cases, comparing the Root Mean Square Deviation (RMSD) between a solution proposed by the algorithm, and a known crystallographic structure of the same complex. Results were deemed conclusive for two of the test cases, but not for the third, which involved large (*i.e.* a larger number of possible solutions) and flexible proteins. However, this work allowed identification of issues related to real-time computation of energy values, as well as search and construction for the flexible structures of proteins. Recent work, carried out by [Ray et al., 2005] and [Ferey et al., 2008], focused on solving these problems through optimized visual rendering using a plugin which interface host molecular visualization an interaction software, such as *Visual Molecular Dynamics (VMD)* and host molecular simulation software, such as *NAMD* or *GROMACS*. The *DockingShop* software program [Lu et al., 2005] currently seems to be the most advanced system for interactive docking. *DockingShop* implements representations of proteins that use principles of robotics, combined to real-time display of essential data for the study of protein-protein interactions (hydrogen bonds, energy, surface complementarity, interpenetration and hydrophobic effects. More recently, a more sophisticated robotics-based approach was proposed to represent peptide chains and molecules [Rossi et al., 2007].

3.3 Main focuses of the *CoRSAIRE* project for docking

However, both these approaches remain very complex regarding the quantity of data that must be conveyed to the user, as well as regarding the limitations of classical WIMP interfaces. Currently, the use of VR devices, such as tracking or haptic devices used in VR, allow more direct and natural interactions with objects in a 3D space. They seem relevant for docking tasks, since these consist in evaluating configurations defined by the relative positions and orientations of one protein in relation to another. Secondly, given the large quantity of information required by users to evaluate the “quality” of a configuration, it also seems relevant to supplement the classical visual feedback with audio and haptic sensorimotor channels. Haptic rendering is known to improve the quality of operator interactivity in an immersive environment, as well as his perception of the objects handled [Seeger & Chen, 1997] or data analyzed [Lundin et al., 2005]. Likewise, audio renderings may improve communication of complex information [Barass & Zehner, 2000]. Furthermore, substitutions and redundancy between these channels of communication may have beneficial results on user performance, as long as the choice of modalities is relevant to the task at hand. For example, [Richard et al., 2006] and [Kitagawa et al., 2005] showed that specific audio and visual renderings can effectively convey information that is usually haptic.

The main focus of the *CoRSAIRE* project for protein-protein docking is to design a new methodology in that field based on advanced interaction and rendering that VR technologies may offer. Complementarily to other works on docking, we are especially studying multi-sensorimotor rendering during an interactive docking task. Several previous works have explored the possibilities afforded by this multimodal feedback for the docking task. However these projects mainly rely on haptic or audio channels, but rarely the combination of both with visual rendering. Haptics-centred projects mainly deal with the generation of force feedback to guide biochemist toward the best solutions. That is the case of *GROPE* [Frederick P. Brooks et al., 1990], *IVPS* [Maciejewski et al., 2005] and *SenSitus* [Wriggers & Birmanns, 2003], a plug-in for *VMD* that allows users to explore proteins with a haptic arm. These projects aim to provide simultaneous renderings of multiple kinds of volumetric data, and to test various paradigms for human-computer interaction. These software platforms allow users to move proteins while receiving haptic feedback regarding collisions between molecules. Evaluations are based on the identification of docking errors as well as the execution time for various paradigms (*e.g.* hard vs. soft docking, monoscopic vs stereoscopic visualization, etc.). Stereoscopic rendering allows reduction of error rates, as well as of the execution time for the various paradigms, whereas haptic feedback lengthens manipulation times due to sequential, local strategies of protein exploration. Other types of haptic feedback, such as vibrotactile feedback, are seldom used. However regarding audio feedback, existing projects aim to provide to users some clues regarding molecular properties (*e.g.* protein binding sites, surface complementarity, etc.) using *earcons* (auditory icons) or data sonification [Garcia-Ruiz & Gutierrez-Pulido, 2006]. Various works concern the sonification of sequential data such as composition of DNA strands [Hart & Read, 2004] or of surface complementarity between proteins, based on computing the standard deviation of minimal distances between pairs of atoms.

Beyond these purely technical aspects, lack of analysis of user needs in the initial stages of such design projects causes gaps between their expected (natural interaction with generated representations of proteins, allowing quick and easy docking of protein complexes) and real results (technically impressive systems that are useless as work

tools). Conversely, our approach relied on early involvement of users in the design process of a docking system, aiming to (1) hasten and ease the development of functional prototypes through improved decision making; (2) improve acceptance of such systems by users of these innovative interfaces; (3) suggest new, unexplored avenues for research and innovation in docking [Anastassova et al., 2007]. Furthermore, current tools for docking put the user in a position of observing and controlling computer-generated solutions, rather than playing an active part in the process. In other words, the scientist is not truly part of the system and cannot tap into his expertise during the search for solutions. The second original aspect of our approach thus aims to involve the scientist in the docking “loop”.

For these reasons, before designing the CoRSAIRe approach for docking (section 4.3) and implementing its concepts into our immersive and multimodal application (section 5), we performed a set of ergonomic studies. So we describe below, the methodology we followed (section 4.1), and the model we obtained about the docking task as it is carried out today, in order to identify advantages and drawbacks of the existing computer-based and VR solutions in that field (section 4.2). These results allowed us to propose a set of basic principles to specify multimodal VR user interface for docking of proteins, which would be likely to allow users to reach the objectives outlined.

4 An initial step for design: clarifying field practices and users needs

4.1 Method

Our study first focused on analyzing the use of existing docking tools, designed for desktop interfaces, as well as the impact, as foreseen by users, of the introduction of VR technology and multimodal interfaces in this task. Four researchers in bioinformatics, aged 28 to 50 years ($M=36$ years, $SD=9.95$) took part in this investigation. Firstly, we carried out, recorded and made *verbatim* transcriptions of four interviews in the workplace. These interviews were anonymous and confidential. They were structured according to a guide which covered various questions ranging from the types of docking problems subjects were confronted to in their line of work, which software was used to solve these, as well as how subjects might envision working in a multimodal virtual environment (MVE). A cognitive discursive analysis [Ghiglione et al., 1998] of the interview corpora was carried out using the program *Tropes* developed by *Acetic Software*. Secondly, three work sessions were videotaped and analyzed, allowing us to view the use of three standard software programs used in docking: *ICM-Disco* [Fernandez-Recio et al., 2003], *ClusPro* [Comeau et al., 2004], and *Hex* [Ritchie, 2003]. Video data of these sessions served as a basis for task analysis. Verbal protocols were collected throughout these sessions, and analyzed using explicit verbalizations of a task as a unit for analysis. Coding the actions and verbalizations collected in this way, allowed us to elaborate two distinct resources for designers. The first, a task model in the form of a hierarchical decomposition using Hierarchical Task Analysis (HTA) methodology [Annett, 2003]; the second, an illustrated overview of the course of a session in the form of a story-board describing the tasks carried out in a standard docking problem: docking of the barnase-barstar complex.

4.2 Results

Results of the analysis allowed us to construct three elements relevant to the design of a MVE for molecular docking: (1) a model of the tasks carried out; (2) a model of user needs and (3) a set of principles to guide design choices in the allotment of information to various display modalities and in designing the associated display and interaction techniques.

4.2.1 A goal-oriented hierarchical model of current task and its anticipated changes with multimodal VR assistance

The HTA task tree (see figure 4) describes the current way of carrying out the docking task as a hierarchy of goals and subgoals. This model has been extracted by combining results from the interviews and observations. In short, the first stage is to generate a large number of potential models of the protein-protein complex. As an input to this task, files are used that describe the 3D structure of proteins in their unbound form, as housed on the *PDB* server. The second stage is the execution of an automatic algorithm to explore the conformational space associated to binding energies (scores) in order to eliminate physically impossible solutions. Clustering then allows grouping of remaining solutions according to similarity. Complex clusters thus produced are then classified according to their score. Finally, the scientist selects a small number of these solutions as candidates for comparative experimental validation. Based on this initial model, we have formalised an hypothetical model of how the task could evolve with multimodal VR in order to provide experts with an improved support to their activities and informational needs related to docking (see figure 5). A more detail account of arguments and bases to this anticipation is given afterwards.

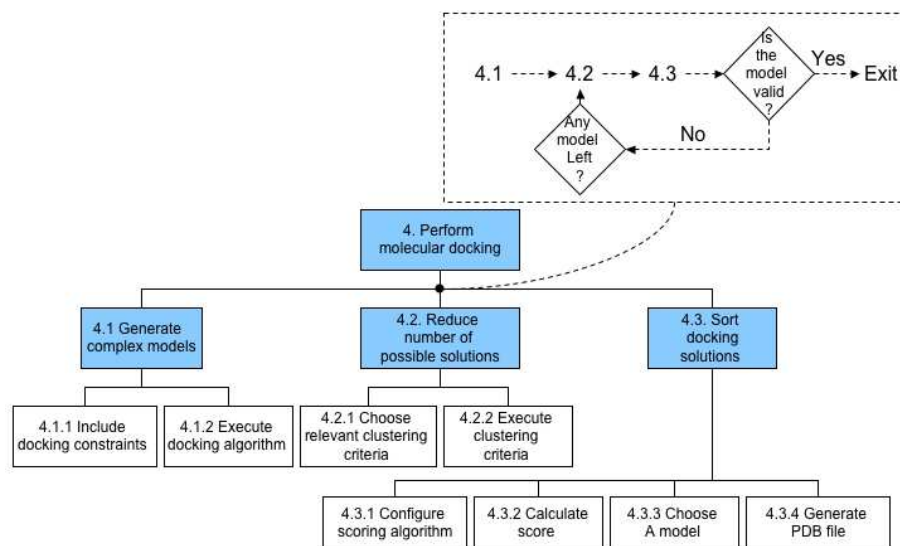


Fig. 4 A partial view of the HTA task tree for molecular docking, and corresponding task plan (inset).

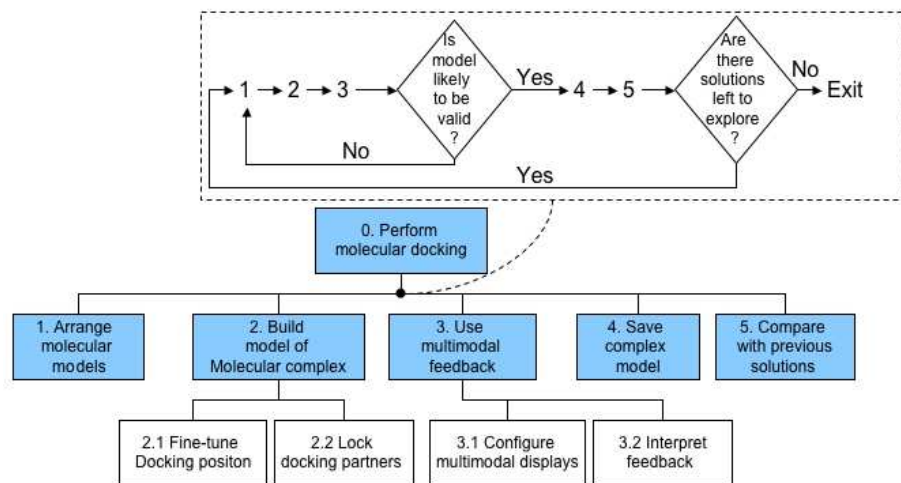


Fig. 5 A proposal for molecular docking as it might be carried out in a multimodal context, and corresponding task plan (inset).

4.2.2 Needs in terms of information display

Cognitive Discursive Analysis highlighted, through balancing and geometrical proximity between central referents in scientists discourse, four basic informational needs in docking. Three of these relate to the properties of the molecules studied. Their interaction is viewed as a central aspect of searching and adjusting for the optimal configuration:

- Topological complementarity of molecules (see section 5.2.1);
- Energy characteristics (see section 5.2.2);
- The fourth need relates to existing knowledge of the molecular interface under study, when such knowledge is available. It consists in a list of amino acids, termed “hotspots”, which the biologist knows to be involved in protein-protein interaction, and therefore to be present at the interface.

One should note that the statistical weight of the score function in the subjects discourse was zero. This suggests that the score function is not viewed as a central tool for docking, but rather as one of several means to confirm hypotheses. This is congruent with the theory underlying scoring functions. Indeed, scores are assessed based on equations physical and chemical phenomena in order to best reflect binding energy. However, in most cases, scoring functions do not allow to discriminate one single complex amongst all generated complexes. Interviews have shown that docking entails joint study of these molecular properties, based on existing knowledge of the molecular interface. Observations of work sessions also showed that scientists spend on average 42% of session time consulting external data sources such as the PubMed database or work notes, in order to identify potential hotspots and construct this knowledge prior to docking.

4.2.3 Restricting the design space for multimodal information distribution by considering usability principles and user task characteristics

The term “modal allocation” [André, 2000] refers to the specific use of one or more sensory modalities to display an information. It is preferable for users to use optimal modal allocation considering both technical (*e.g.* VR-related), task (*e.g.* characteristics of information relevant to scientists) and operator-related constraints (*e.g.* characteristics of perception, of expertise, etc.).

Proposing principles for modal allocation implies weighing specific information-modality associations. One problem is that there exist no *a priori* ergonomic specifications regarding modal allocation for information crucial to such types of tasks. Indeed, although research on the design of human-machine interfaces started several decades ago, few works offer robust principles for choosing a specific modality to display specific types of information. Furthermore, little work focuses on the design of Human Computer Interaction (HCI) for data exploration tasks. It is, however, possible to formulate hypotheses based on the one hand on the ways in which operators process information, and on the other hand (and even more strongly) on the needs highlighted by task analysis. In the case of molecular docking, such analysis has shown the need to display one to four sources of information simultaneously according to what stage the task is at, and to what prior knowledge is available: molecular topology, hydrophobicity of residues, electrostatic fields, and residues potentially involved in the interface.

Task analysis also allowed identification of the following constraints, related to the docking task:

- Since all four types of information may be used by the docking expert at any one time, the risk of overload needs to be anticipated should this information be presented using the same information channel and similar modes. Therefore, information needs to be spread out over various information channels in a balanced fashion to avoid such an overload, particularly visual overload. Although sonification and haptization are liable to effectively lighten the visual channel and improve task performance, randomizing modal allocation may lead to inferior performance to an interface in which all information would be displayed on a single mode. It is therefore necessary to define criteria to decide allocation schemes which would prove to be acceptable to users;
- One important distinction according to this point of view is between data whose status is perennial and invariant as opposed to that which is dynamic and ephemeral, both these types being instrumental to the docking task. Some information, *e.g.* regarding molecular structure, should be available at all times, never changing in the course of one work session. If one were to choose audio or haptic displays for perennial information, one consequence of the properties of cognitive systems for information processing may be perceptual filtering of this information, *i.e.* its disappearance from the focus of attention. One other consequence may be user discomfort, possibly resulting in de-activation of corresponding display functions when the user views the information as already known and overly invasive. For perennial information, we thus suggest to use the visual modality.
- One other criterion for modal allocation is the semantics of information, specifically its proximity to properties of display modalities. As mentioned earlier, taking into account information such as electrostatic forces or hydrophobic interactions is essential to constructing protein-protein complexes. Electrostatic forces might be

	Visual	Auditive	Haptic
Protein surface	ok		
Collisions	ok	ok	ok
Geometric complementarity	ok	ok	ok
Electrostatic interactions		ok	ok
Global electrostatic energy	ok	ok	ok
Van der waals interactions	ok	ok	ok
Global van der Waals energy	ok	ok	ok
Hydrophobic patches	ok		
Hotspots	ok		
Hotspots surface complementarity	ok	ok	ok

Table 1 Restricting the modal allocation space

displayed using a haptic constraint, since this modality can convey attraction or repulsion in a way reminiscent of everyday experience. Visual and audio modalities should not, however, be excluded on principle. But using them would imply constructing an interpretation scale where attraction and repulsion phenomena would be made apparent, as would the scale’s “neutral point”, *e.g.* using one tone for attraction and another for repulsion).

These various constraints allow us to formulate the following principles for the design of a multimodal application for molecular docking, summed up in table 1:

- Use at least the visual modality to display molecular surfaces and contours, as well as allow manipulation of protein models;
- Simultaneously present all information involved in the computation of energy scores, using a combination of modalities;
- As long as it is possible, remain close to the realm of everyday experience, based on information semantics (*e.g.* use haptic rendering for collisions and electrostatic forces);
- Audio signals may be used to sonify time-dependent variables such as the score, presence of hotspots in the interface, presence of hydrogen bonds at the interface, etc.
- Use a combination of modalities to display as much information as feasible without reaching cognitive overload in the user.

It should be noted, however, that these principles are in part hypotheses which will be validated experimentally in the course of an ongoing research project. These validations will be briefly evoked in section 6.

4.3 Reconciling fully automatic and interactive approaches

According to results presented in the previous section, we designed an new approach for protein docking which aims to use expert knowledge by combining multimodal interaction and rendering and automatic approaches. By allowing the user to interact earlier in the process rather than *a posteriori* or through a single, early operation of parameter-setting, we aim to reduce processing times and false positive results (see figure 6). A first stage of the docking task in an immersive environment, named “reduction of the conformational search space”, allows the user, an expert in docking, to interactively build a protein complex (tasks 0.1 and 0.2 of HTA tree in figure 5) which

are potential docking configurations, using visual, audio and haptic feedback (tasks 0.3). This first stage also allows quick reduction of the search space by relying on the expert's 3D pattern matching skills and specific knowledge in protein-protein docking. At this point, a qualitative analysis by interpreting multimodal feedback (tasks 0.3.2) can be done in order to save these docking configurations (tasks 0.4). A second stage named "automatic filtering of selected configurations" relates to the classic automated docking procedure, but will be restricted to search spaces defined by the docking expert. This stage consists in assessing a finer *score* for those conformations selected in the first stage, as well as submitting those to an automatic stage of refining by energy minimization (tasks 0.2.1). Finally, the third stage is called "exploration of the sorted and filtered docking configurations". In this final stage, intermolecular movements are disabled (tasks 0.2.2), because the goal is to explore and compare (tasks 0.5) the solutions generated in the first stage and sort them based on the results of the second stage in order to extract a very small number of protein complexes selected for experimental validation.

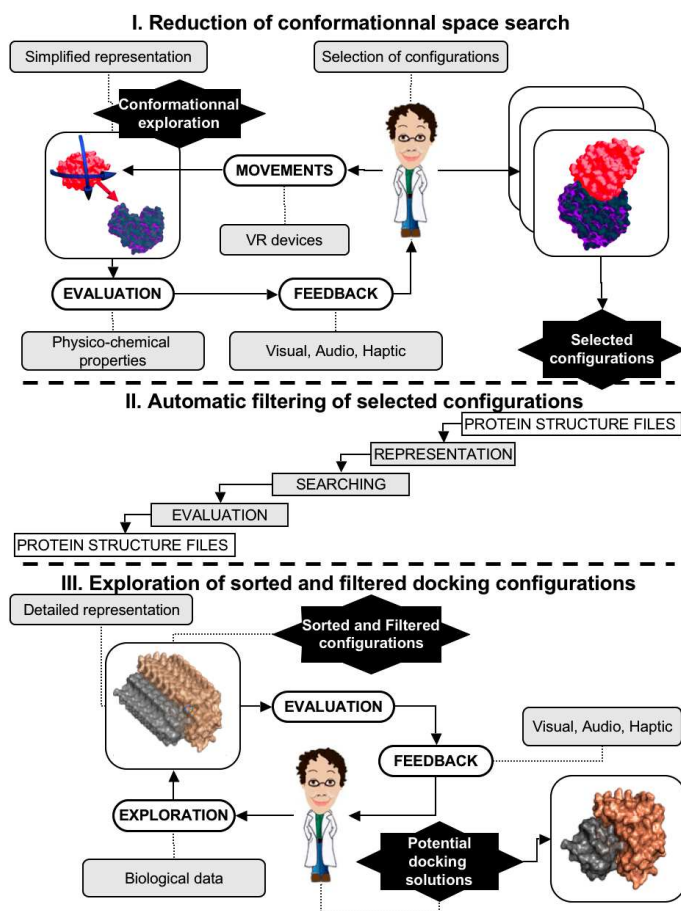


Fig. 6 An hybrid and multimodal approach for docking.

5 Our multimodal and immersive application dedicated to protein docking

5.1 Hardware and Software Architecture

The hardware architecture of our multimodal docking VR environment is based on a CAVE-type device [Cruz-Neira et al., 1992] equipped with a system for overhead projection with an active stereoscopic device for visual immersion. 3D audio feedbacks are generated using *MaX/MSP* and their transmission is carried out using either headphones, or a set of eight loudspeakers spread out over the immersive system. Haptic feedback, finally, is carried out using the *Virtuose* haptic interface, commercialized by *Haption*. Capture of the user’s movements and head movements for active stereoscopy, hand movements to manipulate a 3D mouse is carried out by *ARTrack*, a system of infrared cameras and sensors.

The software architecture of our multimodal docking VR environment (see figure 7) is based on a component programming. Each component can be distributed on network, and is encapsulated into a *Open Sound Control* layer, for managing communication through network between “command” components, “scoring” component, and “rendering” components. Data and events amongst these components are synchronized on network through encapsulation of selected data (the position and orientation of each protein for example) using our own *Open Sound Control* protocol. A “command” component manages distribution of events coming from all VR devices using the software platform *VEserver* [Touraine et al., 2002], except for haptic device, for which we use the *Virtuose* API for managing 6DoF output and force feedback rendering.

The “scoring” component is the physical engine we developed dedicated to protein docking, providing in interactive time biophysical properties of the system, such as energies, forces on atoms or proteins described in section 5.2, and compute visual audio and haptic feedback, sended to the “rendering” components.

The “visual rendering” component is the *Pymol* software, which was adapted for receiving *OSC* messages and interpreting these informations computed by the “scoring” and “command” components, and for displaying these custom informations, in addition to classical molecular representations. The implemented visual feedback are described in section 5.4. The “audio rendering” component is based on *MaX/MSP* software, which already provide *OSC* support, and provide sonification of data coming from “scoring” and “command” components. The implemented audio feedback are detailed in section 5.6. In the “haptic rendering” component, we adapted the *Virtuose* API for receiving *OSC* messages and providing haptic feedback according to of data coming from “scoring” and “command” components. The implemented haptic feedback are described in section 5.6.

Each component has its own representation of proteins and their characteristics, which is well-adapted to the corresponding device constraints and rendering needs. For example, the “visual rendering” component needs all the atom positions, chemical type, and a mesh of protein surface of the two proteins, for providing respectively atom and surface representation, and the “haptic rendering” component needs only mesh of protein surface of the two proteins for computing and rendering rigid body collision.

Finally the “supervisor” component enables or disables feedback and associated computing (feedback control and computing control messages in see figure 7) according to a based-rules system according to dynamic context (scene and computing context messages). A description of the multimodal supervisor is presented in section 5.7.

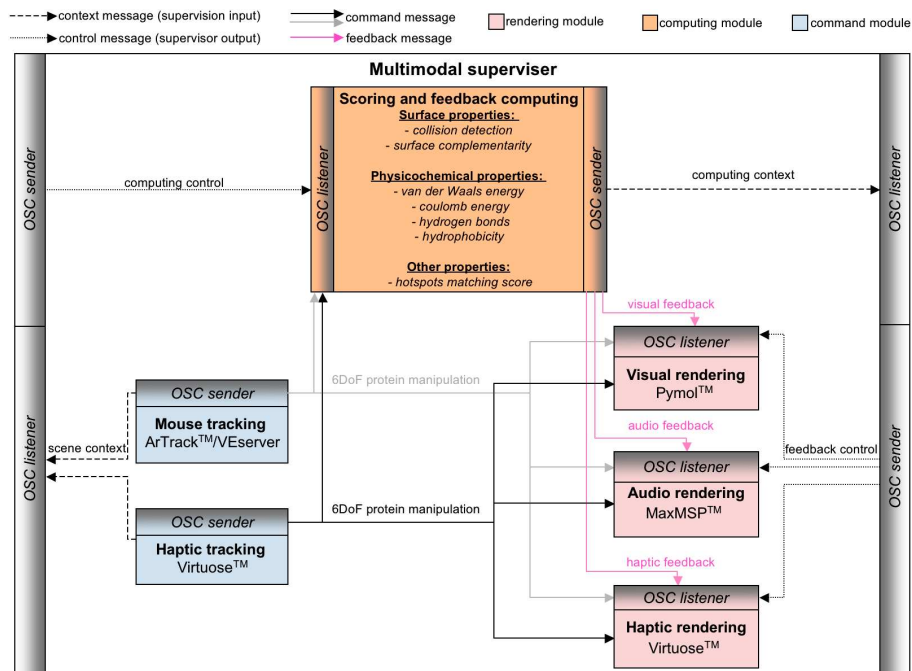


Fig. 7 Software architecture of the CoRSAIRe platform dedicated to multimodal immersive docking

5.2 The “scoring” component : a physical engine dedicated to protein docking

The main constraint we have to overcome in interactive and multimodal docking is to use a docking physical engine, which provides biophysical parameters during the building of a protein complex in interactive time. Moreover, some computation of parameters such as protein surface interpenetration, or electrostatic forces on atoms, have to be very efficient to be haptically rendered. As the automatic docking software presented in section 3.1 do not respond to these constraints, we developed our own physical engine dedicated to protein docking. Protein docking methods are essentially based on two sets of criteria : the geometric/topological criteria, and the biophysical criteria.

5.2.1 Geometry and surface

One of the earliest identified criteria studied in protein-protein interaction is the surface topology of the proteins involved. In most known structures of 3D complexes, partners exhibit good surface complementarity (*e.g.* figure 2B). Studies have also shown that the surface of the protein-protein interface generally covered between 1000 and 2500 angstroms². This criteria allowed the development of first-generation docking software, based solely on shape recognition [Connolly, 1983a] (*i.e.* complementarity of molecular surfaces). This approach is well adapted with rigid proteins docking. We used these geometric/topological criteria in our multimodal immersive environment following two ways:

Collision. For each protein, a surface mesh is computed using *MSMS* before interactive docking [Sanner et al., 1996]. Resolution of this mesh can be set using parameters. The collision detection during interaction then uses the *RAPID* library [Gottschalk et al., 1996], which allows real-time computation of a list of triangles colliding in the two protein surface meshes during docking. This set of triangles can be used to generate feedback based on triangle normals or intersection volume of the two protein surface.

Surface complementarity. It is estimated essentially as a calculation of the variance of the inter-atomic distances on the two proteins surfaces. We use this global surface complementarity score in audio or visual feedback.

5.2.2 Physicochemical properties and energies

However, geometric criteria turned out to be insufficient to predict the structure of a complex. Thus, we include methods based on energy criteria. Protein-protein complexes seem to follow the rule of thumb that the active configuration is that for which the amount of free energy is lowest [Wang et al., 2003]. In order to evaluate free energy between two proteins, we rely on methods of molecular mechanics. To achieve this, atoms are viewed as spheres, and interactions between atoms can be computed using the van der Waals and electrostatic potentials. The free energy for protein-protein interaction can then be approximated by the sum these potentials, which is known as the *score*. In the context of real-time immersive docking, the choice of equations and methods to evaluate a complex’s energy and score is a crucial issue [Wang et al., 2003].

Van der Waals interactions. Van der Waals interactions are an empirical approximation of atomic interactions. The van der Waals force, obtained by constructing a gradient of the potential field, is defined by the Lennard-Jones’s potential equation (equation 1). In this equation, r is the distance between two atoms, σ the interatomic distance for which the potential becomes zero, and ϵ the depth of the potential well (figure 8). ϵ and σ are determined empirically and depend on what pair of atoms is considered. This van der Waals potential includes an attractive component when atoms are bound, and a repulsive component when atoms are too close from each other. It allows preventing two proteins from penetrating into each other during interactive docking, through calculation of interatomic forces at the protein-protein interface.

$$U_{vw}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6\right] \quad (1)$$

These forces apply only to very short distances and only concern surface atoms. As the atom pair distances computing has a quadratic complexity, we apply specific filtering to only keep surface atoms and opposite atoms to the each protein (see figure 9). The resultant translational and rotational components of van der Waals’s forces on each atom are calculated and applied to barycenter protein.

Electrostatic interactions On the contrary to van der Waals interactions, electrostatic interactions operate when “long” distances (about 10 Angström) separate groups of electrically charged atoms. Indeed some amino acids or atoms may present with a positive or negative electric charge, which gives rise to electrostatic phenomena allowing formation of a specific protein-protein complex. Two approaches have been implemented to compute electrostatic phenomena.

We considers interaction between two point charges in vacuum (figure 10), and we use the Coulomb’s law (equation 2) with r being the distance between the barycenters of

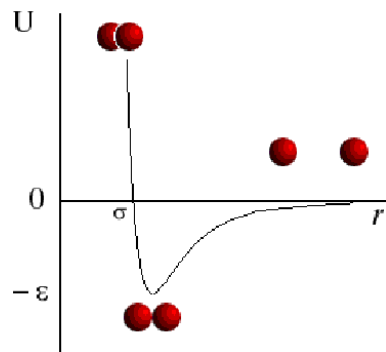


Fig. 8 Lennard-Jones’s potential for van der Waals interactions.

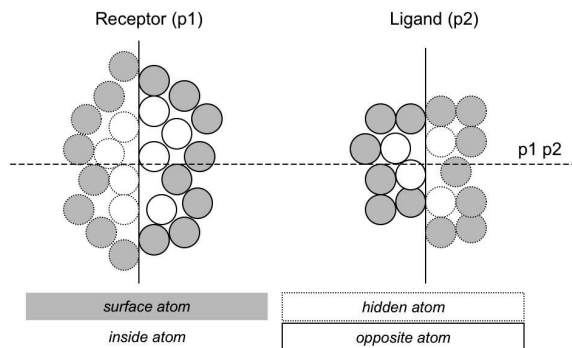


Fig. 9 Dynamic and static atom filtering for optimized van der Waals interactions computing

charges q_1 and q_2 of atoms considered, and ϵ_0 being the constant for the permittivity of a vacuum. This potential can be translated to force (F_{el}) usable for haptic interaction for example (figure 10). This first approach involves calculating the forces to apply to each electrically charged particle considering only pairs of charged particles. This computation has quadratic complexity, because all distance between atoms must be computed, but it remains relevant in the case of medium-sized proteins, since the number of charged particles in a protein is limited in several models.

$$U_{el}(r) = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r} \quad (2)$$

In the second approach (see figure 11, designed for optimisation considerations, the overall field of the electrostatic potential of the target protein (receptor) is computed *a priori* using *APBS*² which allows generation of a 3D grid of electrostatic potential, which can be use as a 3D texture. The gradient of the electrostatic potential allows computation of force field vectors for each point of the grid. Atoms from the ligand protein are then “immersed” in this 3d force field surrounding the receptor. This method allows us to compute electrostatic forces for each atom in linear time, depending on the number of charged atoms in the ligand. In both case, we are able to obtain global

² *Adaptative Poisson-Boltzmann Solver* [Baker et al., 2001]

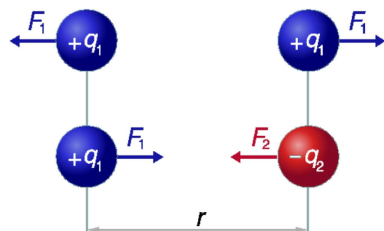


Fig. 10 Resulting forces in the interaction between two point charges, with identical and opposite signs, according to Coulomb's law. Note: $(\frac{1}{4\pi\epsilon_0})$ is also known as K_c or Coulomb's constant.

electrostatic energy and electrostatic force on each atom. These data are then used for visual, haptic or audio feedbacks.

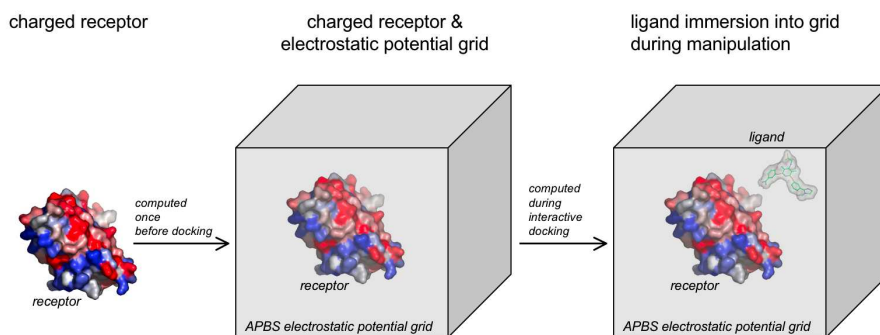


Fig. 11 Ligand immersion in the electrostatic potential grid of the receptor

5.2.3 Other criteria

In order to reach a finer description of protein-protein interactions, other criteria, based on energy, can be taken into account. To geometric/topological criteria, and biophysical criteria, one can add other phenomena which are of utmost importance to protein-protein interaction, such as hydrogen bonds or hydrophobic effects. We only implemented hydrogen bonds in our application.

Hydrogen bonds. Hydrogen bonds (*e.g.* figure 1 in the bottom left corner) are one of the strongest types of interaction in terms of binding energy. On average, there are 5-6 hydrogen bonds per interface protein protein. In our application, when several atoms (nitrogen and oxygen) on the surface of each protein are close enough, under a distance of 3 angstroms, and when their chemical environment is favorable, an hydrogen bonds is created between these atoms. We use the same methods than for van der Waals interactions for filtering atoms on the surface in order to decrease complexity of distance computing between atoms.

5.3 Using haptic device and 3D mouse for bimanual manipulation of proteins

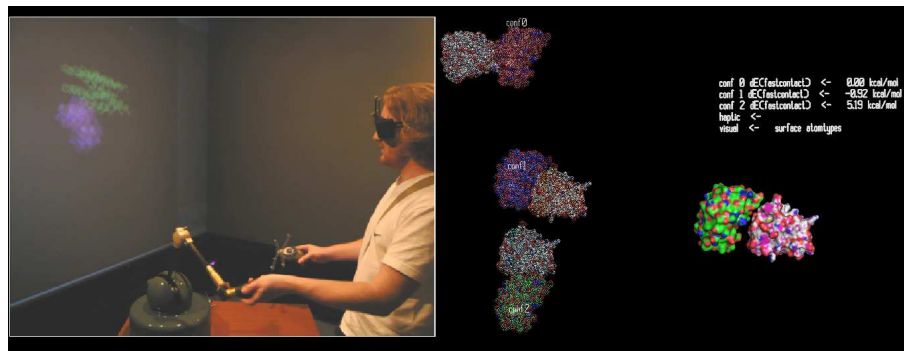


Fig. 12 A user immersed in the docking application (left). On the right, a sample screen capture following selection of three conformations by the user.

In order to manipulate both proteins and attempt to carry out virtual docking, the user may rely on various devices and interaction paradigms (see figure 12). One first paradigm associates the position and orientation of the smaller protein (also called the “ligand”), with the haptic device (6 degrees of freedom, 6DoF), while the second larger protein (also called “receptor”) the orientation of the receptor protein is controlled by a trackball present on the 3D mouse (3 degrees of freedom, 3DoF). To realize this paradigm, the device events are provided by the “command” component and are interpreted by scoring and rendering components (see figure 7). This interaction modality highlights an existing use, since in docking solutions, the receptor protein is static and the exploration concerns position and orientation of the ligand around the receptor. Another paradigm, more in line with natural interaction, consists in manipulating the position and orientation of the ligand and the receptor respectively with the force-feedback device and with the 3D mouse. This modality is symmetrical, in terms of manipulation (*i.e.* commands) but not in terms of renderings, since haptic renderings are only available for one of the two proteins.

Moreover, we used specific method to help user for manipulating tiny objects with haptic device. The basic approach is a direct mapping of the device motion on a molecule, but the interaction aims to act at the atomic level. So, it would be usefull to perform fine tuning operations, especially during rotation movements. Furthermore the physical constraints of the device emerge rapidly during the manipulation. (*i.e.*: workspace of the device). Within the CoRSAIRE project, several haptic paradigms are under investigation to fit better with the accuracy requires by biologist for such manipulation. For instance, *Bubble* method [Dominjon et al., 2005] and *Haptic Hybrid Rotation* one [Dominjon et al., 2006] are centered on the perception of the hardware limitations. When the device is closed to its workspace boundaries, they proposed an elastic force feedback to come back at its neutral position and orientation, and a rate-control of the clutched object. Conversely, when the device is far from these limits, a position control is directly mapped on the clutched object. However, with tiny objects such molecules and atoms, we need mainly a rate-control which, in addition, must be computed with a damping effect. So we are currently investigating if the 6DoF

navigation control damped by a SLERP³ function [Bourdot & Touraine, 2002] may be applied to such accurate haptic manipulations.

5.4 Visual rendering

In order to follow docking expert’s everyday experience, while not re-developing the existing scope of visual representations of proteins, we chose to reuse one of several existing tools for molecular visualization. Our choice was set on the open source software program *Pymol* which allows the use of various visual modalities essential to docking:

The surface representation. The surface representation obtained from various methods depending on the required degree of surface granularity: solvent-accessible surface [Sanner et al., 1996], molecular surface [Connolly, 1983b].

The atomic representations. CPK-type [Corey & Pauling, 1953] for representing atoms as spheres whose radius equals their van der Waals’s radius, stick for representing covalent bonds as tubes or lines, ball-and-stick for representing atomic nuclei as spheres and covalent bonds as tubes or lines, and wireframe for representing atomic bonds as lines.

The secondary structure. The secondary structure representation (α helices and β sheets, see figure 1Bb: ribbon type (a uniform tube follows the carbon chain) or cartoon-type (helices and strands are represented as arrows).

Other criteria were also used in choosing this program, including:

- The fact that expert users were masters in its use;
- The potential to easily integrate specific data obtained a priori in the literature (*e.g.* hotspots) or part of the user’s implicit expertise;
- The existence of a programmable API (Python) allowing speedy integration of new visual or interactive functionalities (under OpenGL);
- A large number of already available scripts and libraries which were relevant to our application.

From these three main types of representations, any combination can be used to visualize a protein. Furthermore, various types of information can be projected onto these representations *e.g.* types of atoms and amino acids on the protein surface; electric charges of these atoms; hydrophobic properties, degree of flexibility, of sequence conservation or *hot spots*. We especially use *Pymol* programmable API, for adding our own visual feedback, for rendering information computed by the docking physical engine:

Energy scores. During the docking task, global energy computed interactively, such as electrostatic energy or van der Waals energy are rendered by text (see 12), when the user need it, or when audio or haptic canal are unavailable.

Hydrogen bonds. In rigid docking, when several atoms (nitrogen and oxygen) on the surface of each protein are close enough, under a distance of 3 angstroms, and when their chemical environment is favorable, an hydrogen bonds is created between these atoms. The potential hydrogen bonds are rendered by lines between atoms.

³ SLERP is shorthand for **spherical linear interpolation** used with **quaternion** representation of 3D rotation.

5.5 Haptic rendering

Currently, there are very few molecular docking systems which include large-scale haptic feedback (force/tactile feedback). This is mainly due to the complexity of computing operations behind the physical engines used for molecular dynamics, which makes it difficult to comply with constraints in terms of refresh rates for real time haptic feedback (from 200 Hz to 1 kHz). This led us to only consider hard docking paradigms, with rigid proteins following a lock-and-key or *LEGO* metaphor. Another difficulty is to render different kind of information: collisions and physicochemical interactions such as van der Waals force and electrostatic force. In order to obtain a coherent haptic feeling, only one type of rendering is provided to the user at a time. However one should note that at the perceptual level, van der Waals's force renderings are similar to surface collision renderings since it prevents interpenetration.

Several haptic renderings were implemented computed with optimized scoring methods described in section 5.2. In all paradigms, we recall that the two proteins was linked to the haptic device and the other to the 3d mouse, which allows to manipulate their position and orientation. The haptic controlled protein is then considered as a big probe with multiple contact points depending on the orientation given by the user.

Van der Waals and electrostatic forces. This rendering is used to provide a biophysical relevant haptic feedback. This haptic rendering of physicochemical interactions consists in feeding the haptic device with the resultant forces computed as described in section 5.2. The forces can be computed and rendered independently or summed up to obtain a total resultant force. The exploration of the receptor by the ligand thus aims at finding stable areas. When the two proteins are in an unstable conformation it renders an unsteady feedback leading the user to keep the ligand at the surface of the receptor and find a better position and orientation. However the complexity of the force fields induces very irregular directional forces affecting the precision of the manipulation. It appears especially with van der Waals interactions because of the non-linearity of the **Lennard-Jones** potential used to model forces.

Collision. Two approaches were explored to render collisions between the two molecules. The first consists in computing a repulsive force. The direction of this force is the opposite of the direction provided and the module is proportional to the number of colliding triangles determined by the RAPID computation as explained in section 5.2. This force can also be weighed by a distance or a volume of interpenetration. Hence the feedback is more relevant but the complexity of the computation induces lower refresh rates which could lead to lags in the feedback. Rather than repulse the two molecules from each other, the second approach, also based on distance computation, aims at preventing collisions locally by modeling contacts points by springs. The method is introduced in [Johnson & Willemsen, 2003] and allows fast computation of local minimum distances based on the geometry of the model as well as resulting force and torque. Interestingly the spring model described can be easily adapted to model atomic clashes, such as van der Waals interactions in our case. Instead of using the complex Lennard-Jones potential to render the resulting force, the interactions are modeled by this more simple spring model with realistic cutoffs (2.5 angstrom). As the atomic distances computation is already optimized to take into account only surface and opposite atoms the refresh rates is sufficient and allows a very precise rendering of the contacts allowing to feel holes and bumps at the surface. Hence the computation speed and coherent feedback constraints are observed keeping a certain biological relevancy. Current researches aim at determining how the size of the proteins affects the

computation time. It will also be interesting to compare this atomic clashes approach with the geometric one which could provide faster computation.

5.6 Audio rendering

Sonification is the use of non-speech audio to convey information. Due to the high temporal resolution and wide bandwidth, the use of auditory stimuli is highly suitable for time-varying parameters (very high temporal definition when compared to other modalities such as video and haptics), concurrent streams (the superposition of multiple audio renderings for various parameters is possible and easily comprehensible if properly designed), and spatial information (lower definition if compared to visual stimuli, but possible over the 360° degree sphere, therefore true full space three-dimensional rendering). Considering the MBI application, the audio channel seems to be well suited for the rendering of time-varying parameters, whether global or distributed locally, which would be difficult to visualize (*i.e.* due to overlapping visual objects, the complex geometrical interface is not visible unless one pries apart the two proteins or alters the ligand display modality).

A large variety of sonification techniques exist and are used in various applications [Walker & Lane, 1994]. One sonification technique is referred to as “parameter mapping” [Hermann & Ritter, 1999], and it is this technique which has been implemented for the current project. Parameter mapping sonification is based on creating a link between the data to be rendered and the parameters of a synthesizer (or of any other device which generates or plays back sound). In this particular sonification typology, three elements need to be carefully considered [Walker & Lane, 2008]:

- The nature of the mapping: which data dimension (*i.e.* temperature, pressure, velocity...) is mapped onto, or represented by, each acoustic parameter (*i.e.* frequency, loudness, tempo...). As an example, for a sonification task the temperature might be linked with the frequency of a sound, therefore as the temperature increases, the frequency of the corresponding sonification increases.
- The polarity of the mapping: to an increase of the data to be sonified, the sonification parameter can decrease or increase. In the case of temperature-frequency mapping, it is common to use an increasing-TO-increasing (up-up) polarity. An alternate example could be the size of an object being mapped to frequency, the polarity would likely be increasing-TO-decreasing such that large objects are linked to low sounds and vice versa.
- The scaling of the mapping: to a determined increase of the data to be sonified, how much does the sonification parameter increase or decrease. One must take into account the possible range of the data, and the percentage of the usable audible range which is to be exploited. Human hearing is more sensitive to small frequency changes at low frequencies, rather than at higher, following an exponential scale. In the case of temperature-frequency mapping the temperature could be linked to the frequency exponentially.

One study was recently performed within this research project regarding the use of sound spatialization, examining the effect of sound spatialization on a specific sonification and sound exploration task [Katz et al., 2008]. Subjects were asked to virtually navigate, using a pointing and tracking device, a two dimensional function mapped onto the surface of a sphere surrounding the user. The data function was sonified with

a modified click/beep sound and the task was simply to find the maximum of the function, the point with the highest frequency beep. The experience was repeated with and without the use of sound spatialization techniques.

In this specific application, sound spatialization is used in two different ways: firstly, for local parameters the sonification is spatialised in the specific position where the parameter is calculated, coherently with the visual or haptic rendering, in order to provide additional information in the proteins coordinate systems (*i.e.* if the task is to sonify the collision between two different atoms on the two proteins, the sonification is spatialised at the position of the collision). Then, multiple concurrent sonifications can be spatially distributed in order to give a better intelligibility of the sonifications themselves (*i.e.* stream segregation, cocktail party effect [Moore, 2003]). In 2007, LIMSI and IRCAM set up a test for the validation of different sonification methods for object manipulation. Within this test, the subject was asked to orient a simplified 3D chemical compound to be the same as that of a given reference. To do this, he/she used an orientation tracking device. Three approaches for data parameter sonifications were tested for improving the speed and accuracy of this manipulation: manipulation speed, angular distance from the reference configuration, and guidance towards the reference position [Arboun, 2007].

Regarding the protein-protein docking task, the following parameters have been selected for the sonification, In the current study, different metaphors have been used for the sonification of the various parameters furnished by the main platform, individually and/or at the same time. Currently, the majority of supplied parameters are global, rather than local. The modules that follow have been specifically designed and developed for this application:

Global score for geometric surface complementarity. This parameter is used to control the variance of a randomly applied pitch to different grains of a granular synthesis process. Granular synthesis has been applied using a spoken word as audio sample (for this particular application, the french word “complementaire” has been recorded and used), repeated cyclically within the granular engine. In this instance, the word is unintelligible if the geometrical complementarity parameter is low, becoming more intelligible as the parameter increases. The rendered audio stream is doubled and associated to each of the two proteins, in preparation for further processing.

Number of “hotspots” at the complex interface. This parameter refers to the list of amino acids present within the current interface region, previously identified using experimental methods as being important actors for protein-protein interaction. Finding hotspots at the protein-protein interface is an important part in judging the quality of solutions. The two audio streams are processed with a low-pass filter with the cutoff frequency controlled by the percentage of proteins hotspots which are situated on the interface region. If none of the hotspots are present on the interface the low-pass filter frequency is set at 200 Hz, making the sound nearly inaudible. The filter’s frequency increases with the number hotspots present at the interface, making the sound clearer and brighter until, in the optimal position, the frequency filtering is completely deactivated. The two audio streams are rendered stereophonically, associating the left and right channels respectively to the first and second protein.

Global electrostatic energy of the complex. This parameter is computed from electrostatic interaction energies between charged particles (*cf.* section 5.2.2). Electrostatic force sonification is performed through the alternation of two sounds, generated using additive synthesis, whose pitch and timbre vary as a function of the global value of this specific force (scalar value). The electrostatic force value is highly

variable, and there is not a direct linear relationship between this parameter and a quality judgement of it being good or bad for the docking condition. The link between the parameter and the quality of its specific value has therefore been traced in a two dimensional Cartesian diagram, with the value of the parameter on the X axis, and the quality (being good or bad) on the Y axis. At a given electrostatic force value, the correspondent value on the Y axis has been sonified with the method previously described. For good values, the frequencies of the two sounds are coincident, and their spectra are perfectly harmonic, whilst as the value worsens, the two frequencies become more distant, and the spectra more inharmonic.

Collisions. One method employed for atomic collision sonification uses a modulation of the phase of a sinusoidal wave whose parameters (carrier and modulator) are controlled by the global number of collisions. Starting with a continuous 400 Hz sinusoidal wave modulated by a 1 Hz signal, the frequency of the modulation increases as the global collision score gets higher, and with it the number of modulating waves, going from 1 to 4, when the two proteins are completely superposed. A second method developed is based on the individual association of every collision with a broadband noise processed with subtractive synthesis (the result is similar to wind noise). The noise is specifically filtered for every collision, adding a controlled randomization of the filtering parameters, so that every “noise generator” sounds different from the others, and spatialised according to its proper position in space. Both of these sonification methods are based on the principle that the signal produced becomes more and more annoying as the number of collisions increases, encouraging the user to change the position and distance of the proteins in order to reduce the number of collisions, and as such stopping the annoying sound.

Van der Waals force. the sonification is based on the principle of the beatings between two sound frequently close. As with the electrostatic force, for the van der Waals force value there is not a linear relationship between the parameter and a quality judgement (being good or bad). A mapping similar to the one described for the previous sonification method (electrostatic force) has been employed, with the Y axis value being sonified. Two intermittent sinusoidal pulses are played back simultaneously: if the quality value for the van der Waals force is good, then the two waves have the same frequency, whilst as it becomes worse, one of the two pulses reduces in frequency by up to 20 Hz from the other. This processing results in the creation of beatings between the two frequencies. If there are no beatings, then the score can be considered to be good. In contrast, if the beatings becomes more frequent (more rapid beat frequency indicates greater frequency separation between the two pulses) the score is becoming worse.

5.7 Multimodal supervisor

As we have seen, part of the choices regarding multimodal allocation is carried out statically, upstream from the execution of the docking application, based on ergonomic recommendations issued following task analysis. However, generation of renderings (or “instantiation”) must be controlled throughout the application. It must depend on characteristics of the data but also on the context of interaction. This context is defined by the user, the real-world environment, virtual environment, and available devices.

To carry out management of multimodal renderings, we modeled and developed a supervision process [Bouyer, 2007] [Bouyer & Bourdot, 2008]. This process was de-

veloped in a generic way, keeping in mind the necessity for it to be able to adapt to various applications, before adapting it to user needs in molecular docking.

This supervision relies on four main actors (see figure 13):

The real world. The docking expert commands the system using the 3d mouse and the haptic device and perceives multimodal renderings via screens, loudspeakers and the same haptic device. He/she is also tracked using various sensors (on the head, in the hand, etc.)

The application. We model the docking application as a trio of elements. The Virtual Environment (VE) contains a virtual representation of the user as well as of the whole dataset involved in the interaction: proteins, scores, numerical variables, etc. The application also rests on hardware and software architectures. Input and output devices are managed by drivers and renderings are generated by specific engines (for graphics, audio and haptic rendering). Finally, the interaction manager interprets user commands, asks the supervisor to control rendering and orders the architecture to display the agreed rendering.

The supervisor. The supervisor is in charge of controlling multimodal rendering. To do this, it relies on a knowledge base containing rules for multimodal allocation (see ergonomic recommendations) as well as a base of rules for decision making which determine its behavior and a contextual base describing elements liable to impact renderings.

The observer/interpreter. This module allows communication between the docking application and the supervisor, by acting as a translator for all exchanged information. Its second role is to survey in real time the running of the application so as to isolate elements of the context and dynamically put them at the supervisor's disposal. Contextual elements might come from the real world (*e.g.* tracking data) and inform for example on the user's position and his movements, or else from the VE and inform on the spatial organization of the scene (relative positions of both proteins, etc.). Knowledge of the application's rendering abilities is particularly useful: regarding media and modalities available in the software, respective load of sensory channels, etc.

More specifically, the supervision process consists in successive exchanges of messages between the following four actors:

- In simplest cases, the user tells the system that he/she wishes to interact with data from the VE (mostly with proteins). The application may also set off an interaction, *e.g.* render a score for electrostatic complementarity when a parameter exceeds a threshold, or navigate following detection of an interesting hotspot.
- The interaction manager interprets this command then makes a request to the supervisor so that it can determine most relevant display modalities. If designers have foreseen a specific rendering scheme (*e.g.* following ergonomic recommendations) or if users have explicitly defined one in their commands, this scheme is appended to the request.
- The request is translated and propagated by the interpreter, to the supervisor. If no rendering is proposed, the supervisor has to specify it completely. If the request is accompanied by proposals, the supervisor will have to validate, reject, or further specify them. The supervisor's decision is a logical process based on static knowledge, dynamic knowledge, and rules applying to this knowledge. First, static knowledge contains the semantic elements of multimodal rendering. They involve choices for information allocation to the various display modalities. Next, dynamic

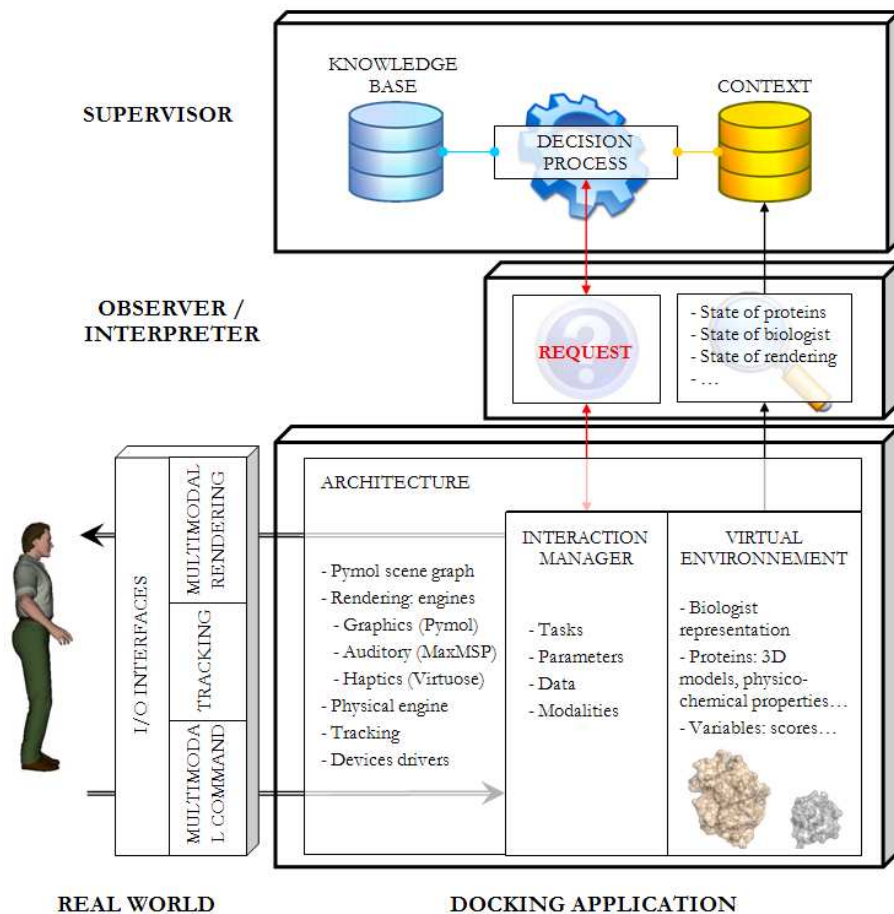


Fig. 13 General architecture for multimodal supervision.

knowledge are contextual elements provided by the observer. Finally, a number of logic rules determine supervisor behavior based on predicates formed by this knowledge.

- The result of the request is processed by the interaction manager. Next, renderings are generated by the appropriate engines and transmitted to the user through relevant media. All contextual information is updated, both in terms of application data and of the supervisor’s contextual database.

6 Evaluations

From the human factors point of view, evaluation of the CoRSAIRe prototype for molecular docking is carried out through an iterative approach of design and evaluation, following the projects a more global user centered design methodology. The two main goals of ongoing evaluations are as follows:

- To evaluate the full virtual environment in terms of usability and usefulness for the targeted population of scientists, the docking task, and the context in which these users carry it out. However, we should point out that the goal here is not so much to evaluate an operational system as to evaluate the relevance of a concept.
- To design and evaluate an interface, notably in terms of designing interaction techniques and multimodal representation of information.

These two goals require very different methods and situations in terms of distance and degree of complexity between devices and users subjected to evaluation on the one hand, and situations of use involving “real” users in a “real-world occupational context on the other hand”. Evaluating a prototype based on its use involves of course relying on realistic scenarios (*i.e.* involving real data and real molecules deemed representative of real problems) and also calling upon representative users. Conversely, when deciding upon choices of modal allocation and evaluating their intelligibility, it is not always compulsory to rely on task experts.

So far, the evaluations we carried out have been formative, *i.e.*; aiming to evaluate throughout product development, the technical and ergonomic quality of prototypes at a given moment, in order to provide designers with feedback and consequently realign design choices. At the end of the project, the approach used is then said to be summative, *i.e.* it aims to assess its quality and properties with reference to external norms and performance criteria. These studies are underway.

7 Conclusion

In this document, we presented an immersive multimodal environment aiming to assist scientists in their study of protein-protein docking phenomena. The first contribution of this work consists of designing new methodology for protein-protein docking 4.3 taking into account advanced interaction and rendering features offering by a Virtual Reality environment. Complementarily to other works on docking described in section 3.1, we adressed the problems relating to multisensorial rendering during the interactive docking task, and the second contribution was to design new haptic feedback 5.5 and audio rendering 5.6 especially dedicated to protein docking, and to provide a method described in section 5.7 for supervising visual, audio and haptic rendering.

Moreover, for designing this new methodology for protein-protein docking and implementing its concepts within our immersive multimodal environment dedicated to protein-protein docking, we followed an iterative user centred design approach. The third contribution of our work was thus an ergonomic study presented in section 4, of existing practices of domain experts in the context of their everyday work. Specifically, our preliminary study mainly aimed to formalize user needs and tasks in order to propose a limited set of design principles and a tool which was adapted to existing working practices. This study mainly aimed to take account of user needs, this set of design principles, and integrate use practices within the immersive multimodal environment.

Our aim was to design an explicitly innovative working environment, yielding new possibilities for use and work practices. Our initial study showed in particular that user participation in the docking task was very limited, since it only involved configuring docking scripts and choosing one result amongst the computer-generated solutions to the studied problem. Indeed we remind that classic approaches to docking provide large numbers of complex configuration based on 3d data describing partner proteins. These algorithms take a long time to produce results, since they test all possible geometric

configurations to dock the two proteins. These configurations are then filtered according to energy and physicochemical criteria. Finally, the scientist selects, in this set of results, a smaller set of possible solutions that can be tested against each other experimentally. Relying on user expertise before applying automatic docking algorithms in a multimodal and immersive context allows the user to use natural abilities for the detection of surface complementarity, as well as prior implicit or literature based knowledge regarding for example the nature of the protein-protein interface, what hotspots are present, etc. We think this process allows significant reduction of the number of configurations to be tested by algorithms used afterwards, and we are currently in intensive evaluation stage to validate this hypothesis, and to evaluate the efficiency of each designed modality for rendering the biophysical parameters during the interactive docking task.

Our approach could be reused in the design of other docking interfaces, integrating factors such as protein flexibility, based on the premise that many docking problems involve flexible partners. Later work will focus on identifying a typology of docking tasks upon which to base a scope of intent. Furthermore, this work should also focus on defining future situations of use of such tools. Indeed, our interactions with future users identified several possible avenues for the use of docking tools, *e.g.* teaching, scientific discovery, collaborative work, etc. All of these situations involve various kinds of constraints and tasks.

The novelty of our approach is that it strives to ensure continuous user participation in the process through direct manipulation of the protein models. In proposing a novel approach in which users are involved both upstream and downstream from automatic docking procedures in a multimodal virtual environment, we hope to maximize the use of his/her expertise. This echoes directly to [Magnani, 153] discussion of assisting scientific reasoning through the use of “epistemic mediators”, *i.e.* external objects which “give rise to new signs, new chances for interpretations, and new interpretations”. In particular, the use of interactive mediators in a multimodal framework would allow widening the perceptual bandwidth [Turk & Robertson, 2000], multiplying the sources of available information *i.e.* topological, electrochemical, a priori knowledge of the protein-protein interface), thereby reducing the risk of false positives which plagues current docking tools.

Acknowledgements This work is currently supported by the *ANR* (the french National Agency for Research) through the *CoRSAIRE* project of *ARA MDMSA* program, and by the *RTRA* (french Thematic Network of Advanced Research) *DIGITEO labs*, through the *SIMCoD* project.

References

- [Anastassova et al., 2007] Anastassova, M., C. Mégard, & J.M Burkhardt 2007. Prototype evaluation and user-needs analysis in the early design of emerging technologies. In 12th International Conference on Human-Computer Interaction, Beijing, China.
- [Anderson & Weng, 1999] Anderson, A., & Z Weng 1999. VRDD: applying virtual reality visualization to protein docking and design. *J. of Mol. Graph. and Mod.*, 17(3):180–186.
- [Andrusier et al., 2007] Andrusier, N., R. Nussinov, & H. J. Wolfson 2007. FireDock: fast interaction refinement in molecular docking. *Proteins*, 69(1):139–59.
- [André, 2000] André, E. 2000. Handbook of natural language processing, chapter The generation of multimedia presentations, pages 305–327.
- [Annett, 2003] Annett, J. 2003. Handbook of Cognitive Task Design, chapter Hierarchical Task Analysis, pages 17–35.

- [Arboun, 2007] Arboun, A. 2007. Evaluation des métaphores de sonification. Master thesis, Ecole Nationale Supérieure Louis Lumière.
- [Baker et al., 2001] Baker, N. A., D. Sept, S. Joseph, M. J. Holst, & J. A. McCammon 2001. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Science*, 98:10037–10041.
- [Barass & Zehner, 2000] Barass, S., & B. Zehner 2000. Responsive Sonification of Well-logs. In *International Conference on Auditory Display*, Atlanta, Georgia, USA.
- [Berman et al., 2000] Berman, H.M., J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, & P. Bourne 2000. The Protein Data Bank. *Nucleic Acids Research*, 1(28):235–242.
- [Borrelli et al., 2005] Borrelli, K.W., A. Vitalis, R. Raul Alcantara, & V Guallar 2005. PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique. *J. Chem. Theory Comput.*, 6(1):1304–1311.
- [Bourdot & Touraine, 2002] Bourdot, P., & D. Touraine 2002. Polyvalent display framework to control virtual navigations by 6DoF tracking. In *Proceedings of IEEE Virtual Reality (IEEE VR 2002, USA, Orlando)*.
- [Bouyer, 2007] Bouyer, G. 2007. Rendu multimodal en Réalité Virtuelle: Supervision des interactions au service de la tâche. Ph.d. thesis, Université Paris XI.
- [Bouyer & Bourdot, 2008] Bouyer, G., & B. Bourdot 2008. Supervision of 3D Multimodal Rendering for Protein-protein Virtual Docking. *Eurographics Symposium on Virtual Environments (EGVE)*, 28:49–56.
- [Comeau et al., 2004] Comeau, S.R., W.D. Gatchell, S. Vajda, & C.J. Camacho 2004. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, 20(1):45–50.
- [Connolly, 1983a] Connolly, M.L. 1983a. Analytical molecular surface calculation. *J. Appl. Crystallogr.*, 16:548–558.
- [Connolly, 1983b] Connolly, M. L. 1983b. Solvent-accessible surfaces of proteins and nucleic acids. *Science*, 221:709–713.
- [Corey & Pauling, 1953] Corey, R.B., & L. Pauling 1953. Molecular Models of Amino Acids, Peptides, and Proteins. *Review of Scientific Instruments*, 24(8):621–627.
- [Cruz-Neira et al., 1992] Cruz-Neira, C., D.J. Sandin, T. A. DeFanti, R. V. Kenyon, & J.C. Hart 1992. The CAVE: Audio Visual Experience Automatic Virtual Environment. *SIGGRAPH*, 35(6):64–72.
- [Dominjon et al., 2005] Dominjon, L., A. Lécuyer, J.M. Burkhardt, G. Andrade-Barroso, & S. Richir 2005. The "Bubble" Technique: Interacting with Large Virtual Environments Using Haptic Devices with Limited Workspace. In *In proceedings of World Haptics Conference (joint Eurohaptics Conference and Haptics Symposium, Pisa, Italy)*.
- [Dominjon et al., 2006] Dominjon, L., A. Lécuyer, J.M. Burkhardt, & S. Richir 2006. Haptic Hybrid Rotations: Overcoming Hardware Rotational Limitations of Force-Feedback Devices. In *In proceedings of IEEE International Conference on Virtual Reality (IEEE VR 2005, Washington, DC, USA)*.
- [Ferey et al., 2008] Ferey, N., O. Delande, G. Grasseau, & M. Baaden 2008. A VR Framework for Interacting with Molecular Simulations. In *ACM-VRST 2008: In proceedings of Virtual Reality Software and Technologies*.
- [Fernandez-Recio et al., 2003] Fernandez-Recio, J., M. Totrov, & R. Abagyan 2003. ICM-DISCO docking by global energy optimization with fully flexible Side-Chains. *Cambridge, MA: Bradford Books / MIT Press*, 1(52):113–117.
- [Frederick P. Brooks et al., 1990] Frederick P. Brooks, Jr., Ming Ouh-Young, James J. Batter, & P. Jerome Kilpatrick 1990. Project GROPE: Haptic displays for scientific visualization. In *SIGGRAPH '90: Proceedings of the 17th annual conference on Computer graphics and interactive techniques*, pages 177–185, New York, NY, USA. ACM.
- [Garcia-Ruiz & Guttierrez-Pulido, 2006] Garcia-Ruiz, M.A., & J.R. Guttierrez-Pulido 2006. An overview of auditory display to assist comprehension of molecular information. *Interacting with Computers*, 18(4):853–868.
- [Ghiglione et al., 1998] Ghiglione, R., A. Landré, M. Bromberg, & P. Molette 1998. L'analyse automatique des contenus.
- [Gottschalk et al., 1996] Gottschalk, S., M.C. Lin, & D. Manocha 1996. OBBTree: A Hierarchical Structure for Rapid Interference Detection. *Computer Graphics*, 30:171–180.
- [Grosdidier, 2007] Grosdidier, A. 2007. Conception d'un logiciel de docking et applications dans la recherche de nouvelles molécules actives. PhD thesis, Université Joseph Fourier Grenoble 1.

- [Hart & Read, 2004] Hart, T.N., & R.J Read 2004. A multiple-start Monte Carlo docking method. *Proteins: Structure, Function, and Genetics*, 13(3):206–222.
- [Hermann & Ritter, 1999] Hermann, T., & H. Ritter 1999. Listen to your Data: Model-Based Sonification for Data Analysis, pages 189–194.
- [Hess et al., 2008] Hess, B., D. Kutzner, C. Vanderspoel, & E. Lindahl 2008. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.*, 4(3):435–447.
- [Hinsen, 2000] Hinsen, K. 2000. The Molecular Modeling Toolkit: A New Approach to Molecular Simulation. *J. Comp. Chem.*, 21:79–85.
- [Johnson & Willemsen, 2003] Johnson, D.E., & P Willemsen 2003. Six Degree-of-Freedom Haptic Rendering of Complex Polygonal Models. In In proceedings of the 11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS’03, Washington, DC, USA).
- [Katz et al., 2008] Katz, FG B, E. Rio, L Picinali, & O. Warusfel 2008. The effect of spatialization in a data sonification exploration tasks. In ICAD 2008: In proceedings of the International Conference on Auditory Display.
- [Kitagawa et al., 2005] Kitagawa, M., D. Dokko, A. Okamura, & D. Yuh 2005. Effect of sensory substitution on suture-manipulation forces for robotic surgical systems. *Journal of Thoracic and Cardiovascular Surgery*, 129(1):151–158.
- [Levine et al., 1997] Levine, D., M. Facello, P. Hallstrom, G. Reeder, B. Walenz, & F. Stevens 1997. Stalk: an interactive system for virtual molecular docking. In *Proceedings of IEEE Computational Science and Engineering*, 4(2):55–65.
- [Lu et al., 2005] Lu, T.-C., J.H. Ding, & S.N. Crivelli Aug 2005. DockingShop: a tool for interactive protein docking. *Computational Systems Bioinformatics Conference, 2005. Workshops and Poster Abstracts. IEEE*, pages 271–272.
- [Lundin et al., 2005] Lundin, K. E., M. Sillen, M. D. Cooper, & A. Ynnerman 2005. Haptic visualization of computational fluid dynamics data using reactive forces. In Erbacher, Robert F., Jonathan C. Roberts, Matti T. Gröhn, & Katy. Börner (eds), *Visualization and Data Analysis 2005. Proceedings of the SPIE, Volume 5669*, pp. 31-41 (2005)., volume 5669 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 31–41.
- [Maciejewski et al., 2005] Maciejewski, S., R. andChoi, D.S. Ebert, & H.Z. Tan 2005. Multi-Modal Perceptualization of Volumetric Data and Its Application to Molecular Docking. In *First Joint Eurohaptics Conference and Symposium on Haptic interfaces For Virtual Environment and Teleoperator Systems*.
- [Magnani, 153] Magnani, Lorenzo 2005 (153). An Abductive Theory of Scientific Reasoning. *Semiotica*, 153(1-4):261–286.
- [Moore, 2003] Moore, B. C. J. 2003. *An Introduction to the Psychology of Hearing*. Academic Press, Londond, UK.
- [Pipe, 2008] Pipe, SW May 2008. Recombinant clotting factors. *Thromb Haemost*, 99(5):840–50.
- [Ray et al., 2005] Ray, N., J.C. Cavin, X.and Paul, & B. Maigret 2005. Intersurf: dynamic interface between proteins. *Journal of Molecular Graphics and Modelling*, 23(4):347–354.
- [Richard et al., 2006] Richard, P., D. Chamaret, F.-X Inglese, P. Lucidarme, & J.-L Ferrier 2006. Human-Scale Haptic Virtual Environment for Product Design: Effect of Sensory Substitution. *The International Journal of Virtual Reality*, 5(2):37–44.
- [Ritchie, 2003] Ritchie, D.W. 2003. Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2. *Proteins*, 52(1):98–106.
- [Rossi et al., 2007] Rossi, R., M. Isorce, S. Morin, J. Flocard, K. Arumugam, S. Crouzy, M. Vivaudou, & S. Redon Jul 2007. Adaptive torsion-angle quasi-statics: a general simulation method with applications to protein structure analysis and design. *Bioinformatics*, 23(13):i408–17.
- [Sanner et al., 1996] Sanner, M., A. Olson, & J.-C. Spohner 1996. Reduced Surface: An Efficient Way to Compute Molecular Surfaces. *Biopolymers*, 38:305–320.
- [Seeger & Chen, 1997] Seeger, A., & J. Chen 1997. Controlling Force Feedback Over a Network. In *The Second PHANToM User’s Group Workshop*, Cambridge, Massachusetts, USA, MIT.
- [Touraine et al., 2002] Touraine, D., P. Bourdot, Y. Bellik, & L. Bolot 2002. A Framework to Manage Multimodal Fusion of Events for Advanced Interactions within Virtual Environments. In *International EUROGRAPHICS Workshop on Virtual Environment*, (EGVE 2002), Barcelona, Spain.

- [Turk & Robertson, 2000] Turk, M., & G. Robertson 2000. Perceptual user interfaces (introduction). *Commun. of ACM*, 43(3):32–34.
- [Villoutreix et al., 2008] Villoutreix, B.O., K. Bastard, O. Sperandio, R. Fahraeus, J.L. Poyet, F. Calvo, B. Deprez, & M.A. Miteva Apr 2008. In silico-in vitro screening of protein-protein interactions: towards the next generation of therapeutics. *Curr Pharm Biotechnol*, 9(2):103–22.
- [Walker & Lane, 1994] Walker, B. N., & D. M. Lane 1994. *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Westview Press.
- [Walker & Lane, 2008] Walker, B. N., & D. M. Lane 2008. Sonification mappings database on the web. In *ICAD 2001: In proceedings of the 2001 International Conference on Auditory Display*.
- [Wang et al., 2003] Wang, R., Y. Lu, & S. Wang 2003. Comparative Evaluation of 11 Scoring Functions for Molecular Docking. *Journal of medicinal chemistry*, 46:2287–2303.
- [Wriggers & Birmanns, 2003] Wriggers, W., & S Birmanns 2003. Interactive fitting augmented by force-feedback and virtual reality. *J. cell biol.*, 144:123–131.
- [Zacharias, 2005] Zacharias, M. 2005. ATTRACT: protein-protein docking in CAPRI using a reduced protein model. *Proteins*, 60(2):252–6.