

Thèse en Informatique

Jean SIMARD

Interactions haptiques collaboratives pour la manipulation moléculaire

École Doctorale d'Informatique de Paris Sud

Thèse soutenue le 1^{er} décembre 2011 en présence de

Martin DUPONT (rapporteur) Directeur de recherche au CNRS-LIMSI
Martin DUPOND (examinateur) Directeur de recherche au CNRS-LIMSI

Table des matières

Table des matières	iii
Table des figures	v
Liste des tableaux	vii
I Le sujet	1
1 Introduction	3
II Étude du travail collaboratif	5
2 La recherche collaborative	7
2.1 Présentation	7
2.1.1 Objectifs	7
2.1.2 Hypothèses	8
2.2 Dispositif expérimental et matériel	8
2.3 Méthode	9
2.3.1 Sujets	9
2.3.2 Variables	9
2.3.3 Tâche	13
2.3.4 Procédure	14
2.4 Résultats	14
3 La manipulation collaborative	17

Table des matières

4	Les dynamiques de groupe	19
III	Propositions pour le travail collaboratif	21
5	Travail collaboratif assisté par haptique	23
IV	Synthèse	25
6	Conclusion et perspectives	27
	Glossary	29
	Acronyms	31
A	Shaddock – Collaborative Virtual Environment for Molecular Design	33

Table des figures

2.1	Illustration du dispositif expérimental	9
2.2	Répartition des résidus sur les molécules	12
2.3	Schéma des phases de la communication verbale	13

Table des figures

Liste des tableaux

2.1	Liste des résidus recherchés	11
2.2	Paramètres de complexité des résidus – Carbone en <i>cyan</i> , Azote en <i>bleu</i> , Oxygène en <i>rouge</i> et Soufre en <i>jaune</i>	11
2.3	Synthèse de la procédure expérimentale	15

Liste des tableaux

Liste des *À faire*

(Section 2.3.4) Compléter les hypothèses	14
--	----

Liste des À faire

Première partie

Le sujet

Chapitre 1

Introduction

Deuxième partie

Étude du travail collaboratif

Chapitre 2

La recherche collaborative

2.1 Présentation

2.1.1 Objectifs

Le contexte de l'expérimentation est l'amarrage moléculaire plus communément nommé *docking* moléculaire. Ce processus implique une analyse et une manipulation complexe reposant sur plusieurs expertises. Il est basé sur une décomposition en trois niveaux de modélisation, traités du niveau le plus grossier au niveau le plus fin :

Niveau inter-moléculaire Cette déformation au niveau macro-moléculaire applique des transformations de grande amplitude sur chaque molécule. L'objectif est de trouver la meilleure concordance entre les molécule en terme de position et d'orientation.

Niveau intra-moléculaire Cette déformation au niveau moléculaire fait suite à la déformation inter-moléculaire. L'amarrage de ces deux molécules (ou plus) introduit de nombreuses interfaces qui doivent être optimisées en fonction de critères variés (la complémentarité des surfaces, les forces électrostatiques, les forces de VAN DER WAALS [MÜLLER 1994] , *etc.*).

Niveau atomique Cette déformation très fine va chercher à optimiser la position des atomes au niveau de l'interface. L'intérêt de cette étape sera portée sur plusieurs types d'interaction (les ponts hydrogènes, les zones hydrophobiques et hydrophylles, les ponts salins, *etc.*).

Pour chacun de ces différents niveaux, le processus de manipulation est similaire et peut être séparé en sous-tâches :

Recherche Cette tâche concerne l'identification et la recherche d'une cible (atome, résidu, hélices α , feuillets β , *etc.*) en fonction de critères multiples (articulations, bilan énergétique, régions hydrophobique, *etc.*).

Sélection Une fois la cible trouvée, la tâche consiste à accéder puis à sélectionner la cible par l'intermédiaire d'un périphérique d'entrée (une souris, une interface haptique, *etc.*).

Déformation La tâche consiste à déformer la structure en manipulant la cible précédemment sélectionnée, que ce soit au niveau inter-moléculaire, intra-moléculaire ou atomique. L'objectif inhérent à cette tâche et d'atteindre l'objectif fixé (par exemple, minimiser l'énergie totale du système).

Évaluation Cette dernière partie va évaluer le travail précédemment réalisé en observant différents indicateurs (énergie potentielle, énergie électrostatique, complémentarité des surfaces, *etc.*). En fonction de la synthèse des résultats de cette dernière phase, un nouveau cycle pourra recommencer (recherche, sélection, déformation, évaluation, *etc.*).

Dans cette première expérimentation, nous proposons d'étudier les premières de ces quatre sous-tâches : la *recherche* et la *sélection*. Ces sous-tâches sont critiques car elles sont nécessaires et vont avoir un impact importants sur les sous-tâches suivantes. Elles vont permettre d'identifier les difficultés liées à la complexité de l'environnement moléculaire (les zones critiques, les zones de flexibilité, *etc.*).

2.1.2 Hypothèses

2.2 Dispositif expérimental et matériel

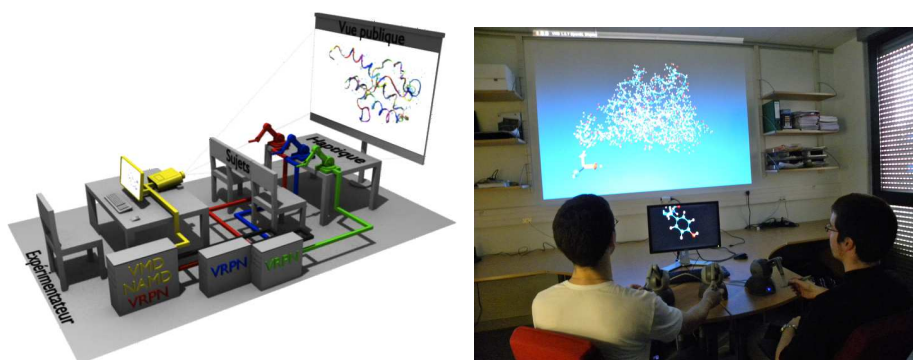
L'Environnement Virtuel Collaboratif (EVC) utilisé est illustré sur la Figure 2.1 page ci-contre. L'EVC propose une visualisation partagée (vue publique) à l'aide d'un vidéoprojecteur sur un grand écran. Le ou les sujets font face à l'écran avec à leur disposition :

- un interface haptique de manipulation *grab* ;
- deux interfaces haptiques de déformation *tug*.

Les sujets ont la possibilité de communiquer entre eux sans restriction. Pour les *monômes*, le sujet peut utiliser chaque outil comme il le souhaite. Pour les *binômes*, chaque sujet se voit attribuer un outil de déformation *tug*. L'outil de manipulation *grab* sera attribuer à un seul des deux sujets après une négociation au sein du *binôme*. Le sujet désigné pour l'outil de manipulation *grab* le sera pour toute la durée de l'expérimentation.

Un micro de bureau est placé en face des deux sujets afin de capter toutes les communications orales. L'enregistrement, réalisé à l'aide du logiciel Audacity, débute à la fin de la phase d'entraînement.

Pour les détails techniques concernant la plate-forme et les outils de manipulation et de déformation, se reporter à l'Annexe A page 33.



(a) Schéma du dispositif expérimental (b) Photographie du dispositif expérimental

Figure 2.1 – Illustration du dispositif expérimental

2.3 Méthode

2.3.1 Sujets

24 sujets (4 femmes et 20 hommes) avec une moyenne d'âge de $\mu = 27.8$ ($\sigma = 7.19$) ont participé à cette expérimentation. Ils ont tous été recrutés au sein du laboratoire Laboratoire pour l'Informatique, la Mécanique et les Sciences de l'Ingénieur (CNRS-LIMSI) et sont chercheurs ou assistants de recherche dans les domaines suivants :

- linguistique et traitement automatique de la parole ;
- réalité virtuelle et système immersifs ;
- audio-acoustique.

Ils ont tous le français comme langue principale. Aucun participant n'a de déficience visuelle (ou corrigée le cas échéant) ni déficience audio.

Chaque participants est complètement naïf concernant les détails de l'expérimentation. Une explication détaillée de la procédure expérimentale leur est donnée au commencement de l'expérimentation mais en omettant l'objectif de l'étude.

2.3.2 Variables

Variables indépendantes

(V_{i1}) **Nombre de sujets** La première variable indépendante est une variable intra-population, c'est-à-dire que tous les sujets seront expérimentés dans toutes les conditions de cette variable. (V_{i1}) possède deux valeurs possibles : « un sujet (*c.f. monôme*) » ou « deux sujets (*c.f. binôme*) ». Les sujets seuls et les sujets en couples ont à leur disposition deux outils haptiques de déformation et un outil

haptique de manipulation. Pour les binômes, seulement un des deux sujets est désigné pour l'utilisation exclusive de la souris 3D. 24 monômes et 12 binômes ont été testés ce qui fait deux fois plus de monômes que de binômes.

(\mathcal{V}_{i2}) **Résidu recherché** La seconde variable indépendante est une variable intra-population. (\mathcal{V}_{i2}) concerne les résidus recherchés qui sont au nombre de 10 répartis à part égale dans deux molécules (voir Table 2.1 page ci-contre). La première molécule est couramment nommée TRP-CAGE [NEIDIGH et al. 2002] et a pour identifiant PDB **1L2Y** sur la *Protein DataBase*¹. La seconde molécule nommée Prion [CHRISTEN et al. 2009] avec l'identifiant PDB **2KFL**. Cinq résidus sont présents sur chaque molécule (voir Figure 2.2 page 12) et chacun présente différents niveaux de complexité (voir Table 2.2 page ci-contre) :

Position La position du résidu peut se trouver sur le pourtour de la molécule, en position *externe* ou à l'intérieur, au milieu de l'amas d'atome (position *interne*). Un résidu en position externe ne nécessite pas de déformer la molécule pour le trouver et l'atteindre contrairement à un résidu en position interne qui sera plus complexe d'accès.

Forme La forme du résidu influe énormément sur la complexité de la recherche. On distingue trois formes différentes :

Chaîne Un enchaînement d'atomes seuls les atomes d'hydrogène sont de part et d'autres de cet enchaînement.

Cercle Une chaîne d'atomes de carbone ou d'azote qui boucle sur elle-même.

Étoile Séries de chaînes d'atomes toutes reliées sur un atome central (la plupart du temps, un atome de carbone).

Couleurs Les atomes sont colorés en fonction de leur nature (rouge pour l'oxygène, blanc pour l'hydrogène, *etc.*). Les atomes qui sont rares seront donc rapidement trouvés grâce à leur couleur différente. Par contre, les atomes nombreux (comme les hydrogènes ou les carbones) seront plus difficiles à filtrer à cause de leur nombre important.

Similarité Certains résidus à chercher sont très similaires à d'autres résidus également présents sur la molécule. De par leur similarité, ils vont mobilier la recherche sur des résidus incorrects.

Variables dépendantes

(\mathcal{V}_{d1}) **Le temps de complétion** Ce temps est le temps total pour réaliser la tâche demandée, c'est-à-dire trouver le résidu et l'extraire de la molécule. Ce temps est divisé en deux phases bien distinctes :

La recherche C'est la phase pendant laquelle les sujets cherchent le résidu. Cette recherche peut être simplement visuelle en orientant et en déplaçant

1. <http://www.pdb.org/>

Table 2.1 – Liste des résidus recherchés**(a)** Résidus sur la molécule TRP-CAGE**(b)** Résidus sur la molécule Prion

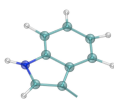

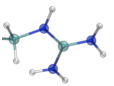
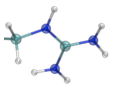
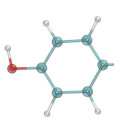
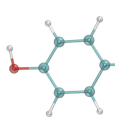
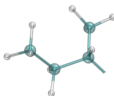
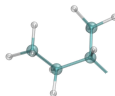
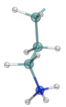
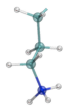
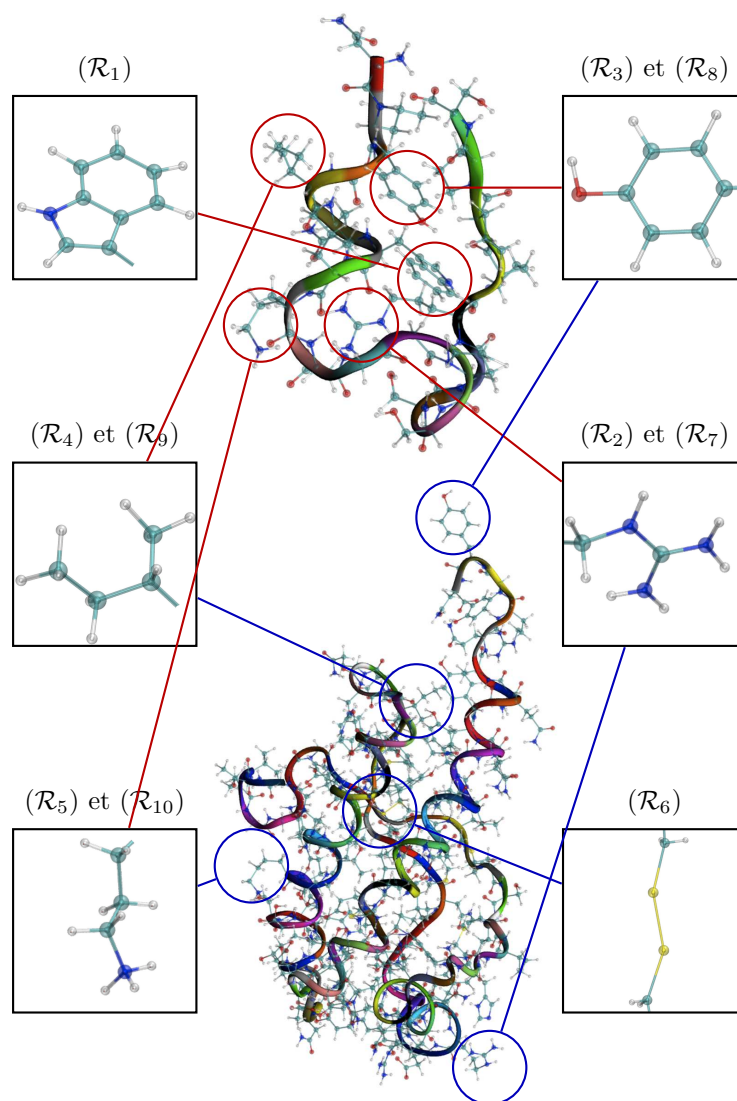
Résidu	Image	Résidu	Image
(\mathcal{R}_1)		(\mathcal{R}_6)	
(\mathcal{R}_2)		(\mathcal{R}_7)	
(\mathcal{R}_3)		(\mathcal{R}_8)	
(\mathcal{R}_4)		(\mathcal{R}_9)	
(\mathcal{R}_5)		(\mathcal{R}_{10})	

Table 2.2 – Paramètres de complexité des résidus – Carbone en cyan, Azote en bleu, Oxygène en rouge et Soufre en jaune

Résidu	Position	Forme	Couleurs	Similarité
(\mathcal{R}_1)	Interne	Cercle	8 C, 1 A	Non
(\mathcal{R}_2)	Interne	Étoile	1 C, 3 A	Non
(\mathcal{R}_3)	Interne	Cercle	6 C, 1 O	Non
(\mathcal{R}_4)	Externe	Chaîne	4 C	Non
(\mathcal{R}_5)	Externe	Chaîne	4 C, 1 A	Non
(\mathcal{R}_6)	Interne	Chaîne	2 C, 2 S	Non
(\mathcal{R}_7)	Externe	Étoile	1 C, 3 A	Non
(\mathcal{R}_8)	Externe	Cercle	6 C, 1 O	Non
(\mathcal{R}_9)	Interne	Chaîne	4 C	Oui
(\mathcal{R}_{10})	Interne	Chaîne	4 C, 1 A	Oui

Figure 2.2 – Répartition des résidus sur les molécules



la molécule mais elle peut aussi amener les sujets à déformer la molécule afin d'explorer les résidu inaccessibles.

La sélection La phase de sélection débute dès l'instant où un des deux sujets a trouvé le résidu. Elle est constitué d'une phase de sélection puis d'une phase d'extraction.

Il n'y a pas de limite de temps pour réaliser la tâche.

(\mathcal{V}_{d2}) **La distance entre les espaces de travail** Cette distance est la distance moyenne entre les deux effecteurs terminaux présents durant l'expérimentation. Cette distance représente donc une distance physique du monde réel, pas une distance virtuelle. Elle est de l'ordre du centimètre.

(\mathcal{V}_{d3}) **Les communications orales** L'enregistrement audio permet de mesurer la quantité de temps de parole pendant chaque tâche de l'expérimentation. Ces mesures discrimine la phase de recherche de la phase de sélection (voir (\mathcal{V}_{d1})) comme indiqué plus précisément sur la Figure 2.3.

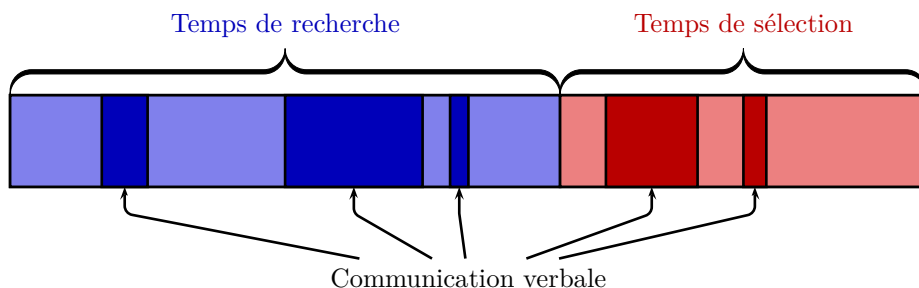


Figure 2.3 – Schéma des phases de la communication verbale

(\mathcal{V}_{d4}) **L'affinité entre les sujets** Le degré d'affinité – concernant uniquement les binômes – est compris entre 1 et 5 selon les critères suivants :

1. Les sujets ne se connaissent pas ;
2. Les sujets travaillent dans la même entreprise, le même laboratoire ;
3. Les sujets travaillent dans la même équipe ;
4. Les sujets travaillent dans le même bureau ;
5. Les sujets sont amis.

2.3.3 Tâche

La tâche proposée est la recherche et la sélection dans un EVC sur des molécules complexes. Les motifs à rechercher dans les structures moléculaires sont des

résidus (voir Table 2.1 page 11). Les sujets possèdent deux outils pour trouver ces motifs :

- ils peuvent explorer la molécule en la déplaçant ou en la tournant à l’aide de l’outil *grab* ;
- ils peuvent déformer la molécule à l’aide de l’outil *tug*.

Une fois le résidu trouvé, les sujets doivent le sélectionner puis l’extraire hors de la sphère virtuelle englobant la molécule.

2.3.4 Procédure

Pour débiter cette expérimentation, les sujets sont confrontés à un exemple sur la molécule TRP-ZIPPER [COCHRAN et al. 2001] ayant pour identifiant PDB 1LE1. Pendant la phase d’entraînement, les outils sont introduits et expliqués un par un. Chaque sujet a la possibilité de tester les outils et peut questionner l’expérimentateur. Dans le cas des binômes, cette phase d’entraînement est également l’occasion de choisir qui, parmi les deux sujets, sera en charge de la manipulation de la molécule à l’aide de l’outil de manipulation *grab*.

Au début de l’expérimentation, l’expérimentateur affiche sur l’écran annexe le motif à chercher. Lorsque les sujets ont trouvé et extrait le motif, l’expérimentateur indique au système la fin de la première tâche. Il affiche à présent le second motif et ainsi de suite pour les dix motifs.

Une fois le groupe passé dans sa première configuration, monôme ou binôme (voir (\mathcal{V}_{i1})), l’opération est répétée pour les autres configurations.

Compléter les hypothèses

Les hypothèses ne sont pas encore écrites!! Il va donc falloir les compléter dans le tableau.

Incomplet – Compléter les hypothèses

Les hypothèses ne sont pas encore écrites!! Il va donc falloir les compléter dans le tableau.

2.4 Résultats

Table 2.3 – Synthèse de la procédure expérimentale

Tâche	Recherche et sélection de motifs	
Hypothèses		
Variable in-dépendantes	(\mathcal{V}_{i1}) Nombre de sujets (\mathcal{V}_{i2}) Résidu à chercher	
Variable dépendantes	(\mathcal{V}_{d1}) Temps de complétion (\mathcal{V}_{d2}) Distance entre les espaces de travail (\mathcal{V}_{d3}) Communication orales (\mathcal{V}_{d4}) Affinités entre les sujets	
Condition (\mathcal{C}_1)	Condition (\mathcal{C}_2)	Condition (\mathcal{C}_3)
Sujet <i>A</i> 10 résidus	Sujet <i>A</i> 10 résidus	Sujet <i>A</i> et <i>B</i> 10 résidus
Sujet <i>B</i> 10 résidus	Sujet <i>A</i> et <i>B</i> 10 résidus	Sujet <i>A</i> 10 résidus
Sujet <i>A</i> et <i>B</i> 10 résidus	Sujet <i>B</i> 10 résidus	Sujet <i>B</i> 10 résidus

Chapitre 3

La manipulation collaborative

Chapitre 4

Les dynamiques de groupe

Troisième partie

Propositions pour le travail collaboratif

Chapitre 5

Travail collaboratif assisté par haptique

Quatrième partie

Synthèse

Chapitre 6

Conclusion et perspectives

Glossary

amarrage moléculaire

Méthode permettant de déterminer l'orientation et la déformation optimale de 2 molécules afin qu'elle s'assemble pour former un complexe stable. 7, 29

binôme

Groupe constitué de 2 personnes. 8, 9, 13, 14

curseur

Élément virtuel associé à un élément physique que le sujet manipule ; il est lié à l'effecteur terminal. 29

docking moléculaire

Voir amarrage moléculaire. 7

effecteur terminal

Élément physique que le sujet manipule ; il est lié au curseur du monde virtuel. 13

monôme

Groupe constitué d'une unique personne. 8, 9, 14

résidu

Groupe d'atomes constituant un des blocs élémentaires d'une molécule. 7, 10, 13, 14

variable dépendante

Facteur mesuré sur une expérimentation (nombre de sélections, trajectoire, *etc.*) ; ces variables sont influencées par les variables indépendantes. 29

variable indépendante

Facteur pouvant varier et être manipuler sur une expérimentation (nombre de participants, tâche, *etc.*) ; ces variables vont avoir une incidence sur les variables dépendantes. 9, 10

variable intra-population

Variables pour lesquelles les sujets sont confrontés à toutes les modalités de la variable. 9, 10

Acronyms

evc

Ensemble logiciel et matériel permettant de faire interagir plusieurs utilisateurs au sein d'un même environnement ; ils jouent un rôle important dans le développement de nouvelles méthodes de travail collaboratives. 8, 13

cnrs-limsi

Unité Propre de Recherche du CNRS (UPR3251) associé aux universités PARIS Sud et Pierre et Marie CURIE. i, 8

Acronyms

Annexe A

Shaddock – Collaborative Virtual Environment for Molecular Design