

5. MODELISATION DU PROTEASOME DE *P. PATENS*

5.1 INTRODUCTION GENERALE.

D'un point de vue académique, il est essentiel de connaître la forme ou la conformation tridimensionnelle d'une protéine pour comprendre son interaction avec d'autres protéines, de façon à mieux comprendre son rôle dans la cellule. Les protéines ont des structures tridimensionnelles très spécifiques qui déterminent leur fonctionnement dans la cellule, et notamment leur activité biologique.

Les méthodes de cristallographie par diffraction aux rayons X et la spectroscopie par résonance magnétique nucléaire permettent de donner une image assez précise des structures des protéines, mais ce sont des techniques que l'on ne peut appliquer à toutes les protéines. De plus, elles sont coûteuses et fastidieuses. Ainsi, la relative difficulté d'élaborer des structures en trois dimensions de protéines peut s'évaluer au fait que sur plus de 200'000 protéines déjà séquencées, nous ne connaissons la structure que d'environ 4'000 d'entre elles. La prédiction de la structure d'une protéine directement à partir de sa séquence en acides aminés éviterait la nécessité de synthétiser ou de purifier de grande quantité de celle-ci pour une détermination de la structure cristalline par rayons X. Pourtant, on ne comprend pas parfaitement la relation qui existe entre la structure tertiaire et la séquence primaire de la protéine. Un certain nombre d'expériences démontre que le nombre de "voies" par lesquelles une protéine peut acquérir sa conformation finale est limité. De plus, la conservation de la structure a largement prédominé sur la conservation de la séquence durant l'évolution des protéines.

A l'heure actuelle, deux méthodes principales de modélisation semblent être couramment utilisées. D'une part, la modélisation par homologie ou comparaison; lorsqu'il y a une relation claire entre la séquence nouvelle en acides aminés et les structures connues de protéines. De tels alignements de séquences peuvent être effectués au moyen de programmes (BLAST, FASTA). D'autre part, la structure tridimensionnelle peut-être prédite à partir des propriétés physico-chimiques des acides aminés qui constituent la protéine.

Ce chapitre décrit la modélisation d'une structure complexe et symétrique sur la base de photographies obtenues en microscopie électronique, en utilisant une nouvelle technique, dite de *cryo-negative staining*. Celle-ci permet d'obtenir un volume de reconstruction de la protéine, fidèle à la structure obtenue par cristallographie R-X. Nous tenons à signaler que ce travail a été effectué en collaboration avec Messieurs Marc Adrian et Eric Larquet du laboratoire d'analyse ultrastructurale de l'Université de Lausanne (LAU), dont le directeur est le professeur Jacques Dubochet. En tant que biologistes, nous ne prétendons pas posséder l'expérience nécessaire ni la formation suffisante pour mener de bout en bout ce type de travail sans l'aide de physiciens. Les résultats présentés dans ce chapitre ne reflètent pas la somme totale des informations obtenues au cours du travail de modélisation. Ils ne font que marquer les étapes principales du processus de base sur lequel la modélisation s'est appuyée. L'ensemble du traitement des images ou données a été effectué à l'aide de programmes informatiques.

5.2 MICROSCOPIE ELECTRONIQUE

La technique de coloration négative dite de *negative staining* est la plus couramment utilisée pour visualiser en microscopie électronique des préparations contenant des particules biologiques. La résolution d'une telle technique se situe au environ de 20 Å. Cette technique rapide, simple et reproductible, donne des images fortement contrastées. Pourtant elle comporte plusieurs désavantages inhérents à la préparation de l'échantillon puisque celui-ci doit être déshydraté avant d'être observé sous le faisceau électronique. Ces désavantages – la déformation des structures, une résolution plus faible, des structures qui ne se trouvent plus dans le contexte physiologique dont elles sont issues (salinité, pH, etc.) – ne donnent pas toujours une image fidèle de l'échantillon observé.

Ces problèmes de déshydratation ont été résolus par le développement au début des années 80' d'une technique de vitrification en milieu aqueux des structures biologiques. Celles-ci sont figées dans de la glace, dite "amorphe", exempte de cristaux, ce qui évite les déformations. Les structures se trouvent en milieu hydraté proche de leur milieu physiologique où la force ionique est respectée. La résolution des images obtenues peut aller jusqu'à 3Å. De plus, les molécules prises dans cette masse de glace amorphe peuvent se positionner de manière aléatoire. L'observation de telles préparations offre la possibilité de visualiser la structure sur toutes ses coutures. Ce qui n'est pas le cas de la méthode par coloration négative où les molécules ne peuvent se placer que dans des positions déterminées suivant la conformation de la protéine. Cette technique dite de *cryo-electron microscopy* est plus délicate et peut parfois présenter un contraste plus faible que la méthode de coloration négative.

Ce travail présente pour la première fois l'application d'une nouvelle technique de microscopie électronique dite de *cryo-negative staining* qui offre les avantages de l'une et de l'autre des techniques présentées précédemment, c'est-à-dire le respect des structures biologiques, dans un environnement hydraté, en présence d'un agent contrastant (molybdate d'ammonium) enrobant les particules figées dans la glace. On peut extraire de telles images beaucoup plus d'informations sur une même particule qu'en coloration négative.

Le protéasome de mousse a donc joué le rôle de cobaye dans ce qui peut constituer ici une sorte de validation de cette nouvelle technique. C'est la première structure modélisée grâce à cette nouvelle technique.

La figure 5.1 présente une photographie obtenue par la technique de *cryo-negative staining*. Trois vues principales se dégagent de cette préparation. La vue de dessus et de côté du protéasome de mousse, ainsi que des disques alpha libres vus de dessus (flèches). Ces anneaux libres sont moins visibles, à l'œil du moins, puisqu'ils n'ont pas la même densité électronique que les anneaux alpha placés dans la structure du protéasome vus de dessus. Avec une telle technique, il s'agit d'obtenir des préparations de protéasome très pures et suffisamment diluées pour obtenir des structures suffisamment distantes les unes des autres.

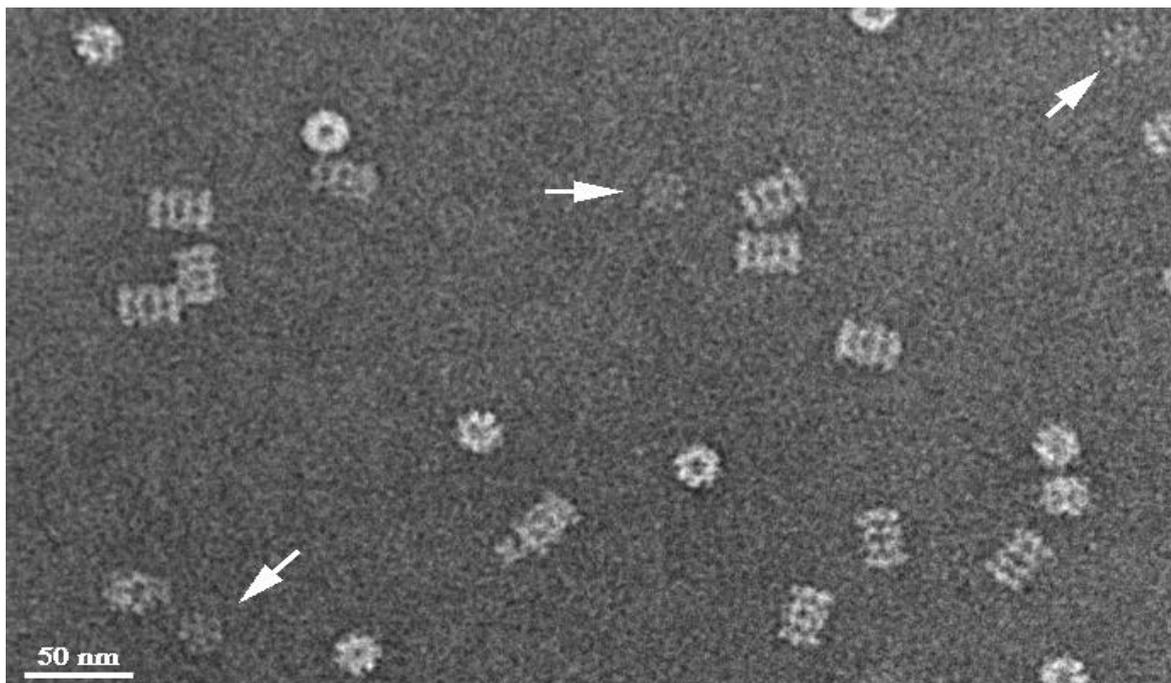


Fig. 5.1. Micrographie obtenue par la méthode dite de " cryo-négative staining " d'une préparation de protéasome de *P. patens*. Outre les structures

Principes généraux de la modélisation

Nous décrivons ici les étapes principales de la modélisation du protéasome de mousse à partir de micrographies. La première partie du travail est appelée analyse bidimensionnelle. Les négatifs des images obtenues en microscopie électronique sont balayés ou "scannés" à haute résolution pour les digitaliser de façon à pouvoir les traiter au moyen de programmes informatiques. Ensuite, il s'agit de sélectionner ou de "fenêtrer" les structures qui les composent. C'est-à-dire de les inscrire dans un cadre. Les images ainsi sélectionnées et numérotées sont groupées dans un des trois répertoires possibles, soit les vues de dessus (fig. 5.2 A), les vues de côtés (fig. 5.3 A), et les vues des anneaux alpha libres (fig. 5.4.A).

Les structures que l'on observe dans chacune de ces images sont disposées de façon aléatoire. A ce stade, le programme informatique peut les traiter en effectuant un "moyennage". Autrement dit, une superposition de toutes les images d'un même répertoire. Pour cela, il s'agit de réorienter toutes les images par rapport à un même référentiel choisi par le programme de façon arbitraire. Les figures 5.2 B, 5.3.B et 5.4.B présentent respectivement pour les trois vues principales, la réorientation des images observées en figures 5.2.A, 5.3 A et 5.4 A. Le programme informatique peut alors procéder à la superposition de ces images pour chacun des répertoires. Les quatre images (de droite à gauche) des figures 5.2 C, 5.3 C et 5.4 C, représentent le résultat de quatre cycles d'affinement successifs de ce "moyennage". L'opération de "moyennage" ne donne, par définition, qu'une information moyenne de la structure bidimensionnelle.

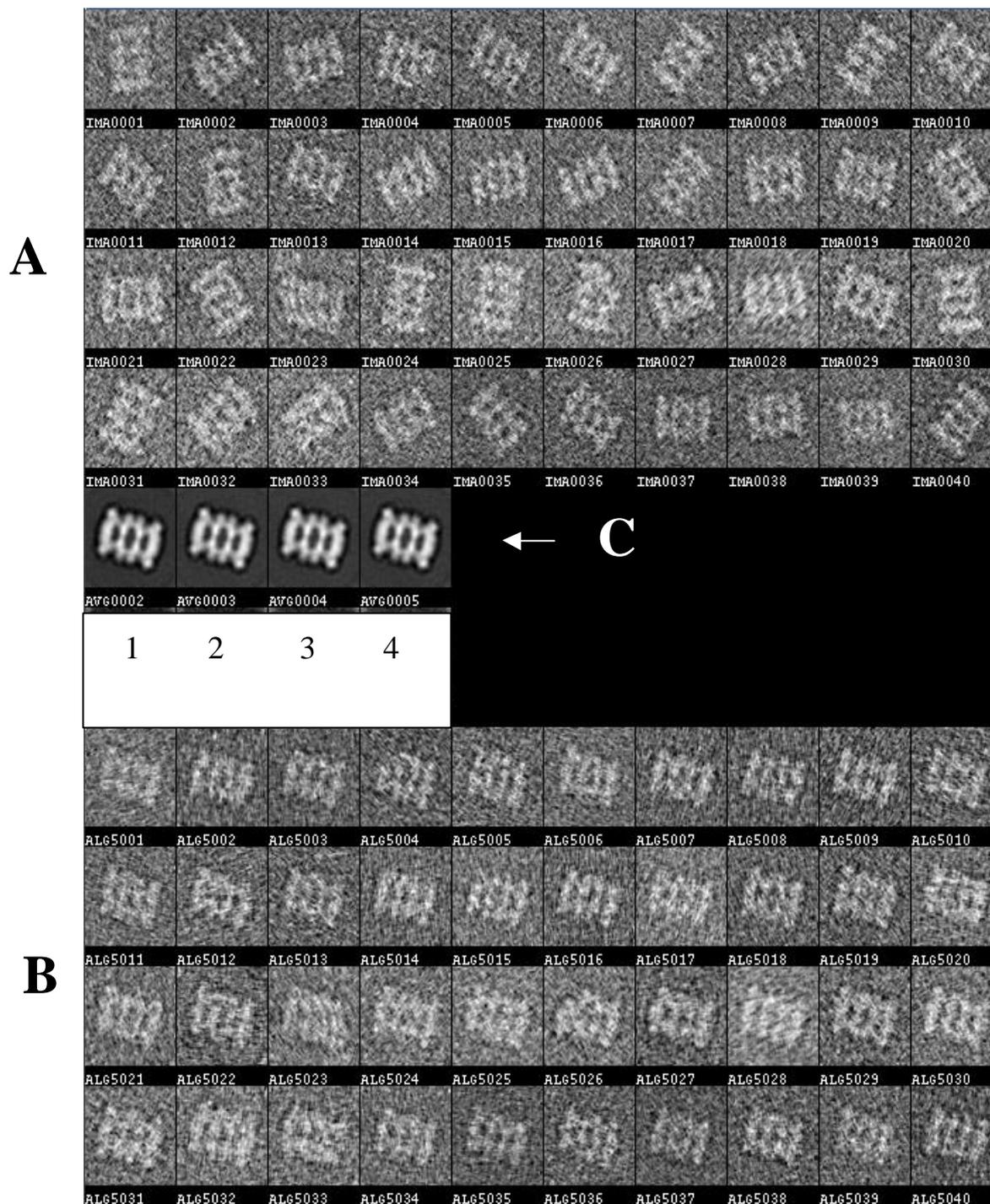


Fig. 5.2. Analyse bidimensionnelle des structures vues de côté du protéasome. **A)** sélection des images appartenant à une même vue. **B)** réorientation des images sélectionnées par rapport à un même référentiel. **C)** images obtenues après superposition ou "moyennage" des images sélectionnées. Chaque image représente un cycle d'affinage supplémentaire. Image 1=1cycle, image 4=4 cycles.

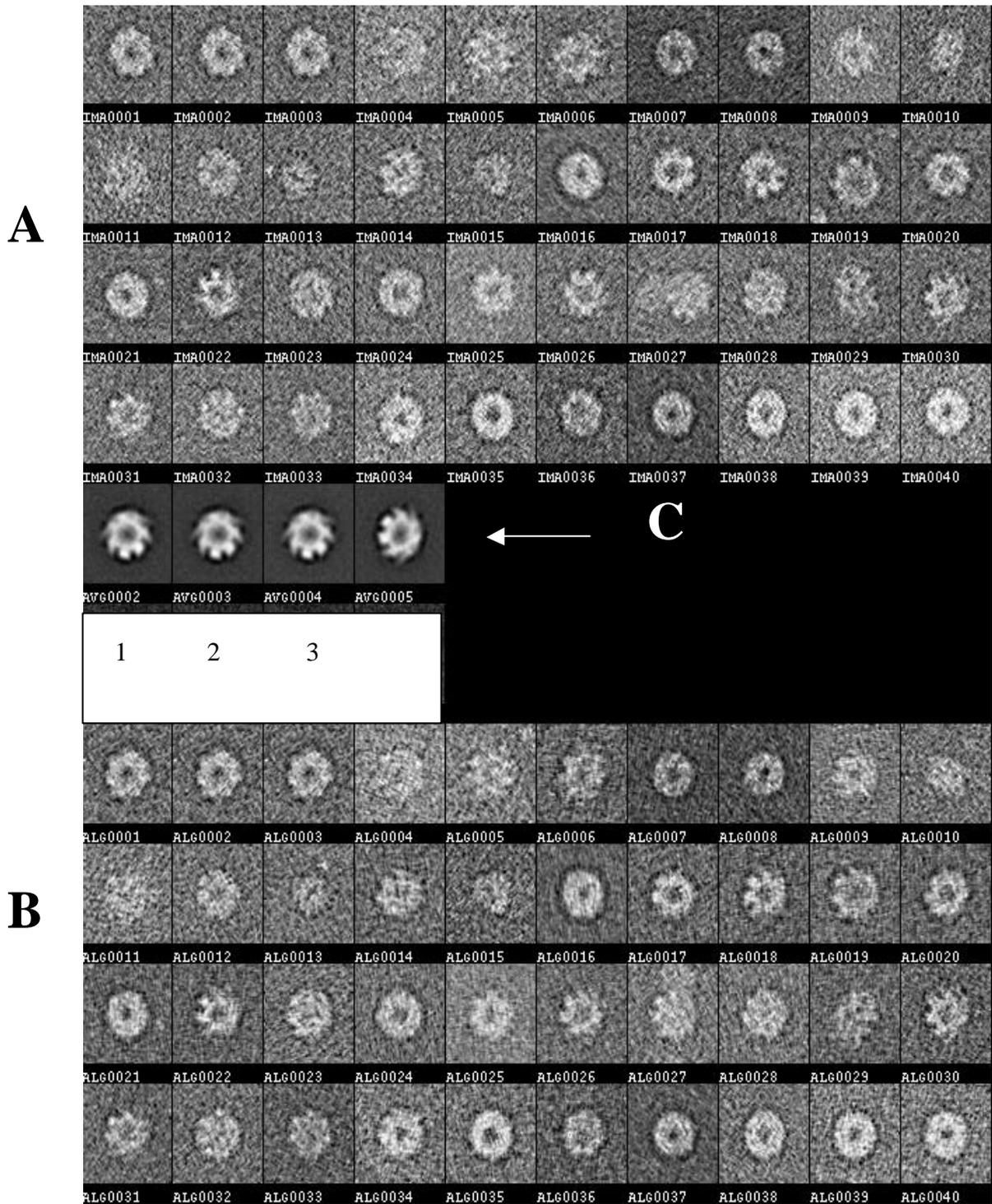


Fig. 5.3. Analyse bidimensionnelle des structures vues de dessus du protéasome. **A)** sélection des images appartenant à une même vue. **B)** réorientation des images sélectionnées par rapport à un même référentiel. **C)** image obtenue après superposition ou "moyennage" des images sélectionnées. Chaque image représente un cycle d'affinage supplémentaire. Image 1=1cycle, image 4=4 cycles.

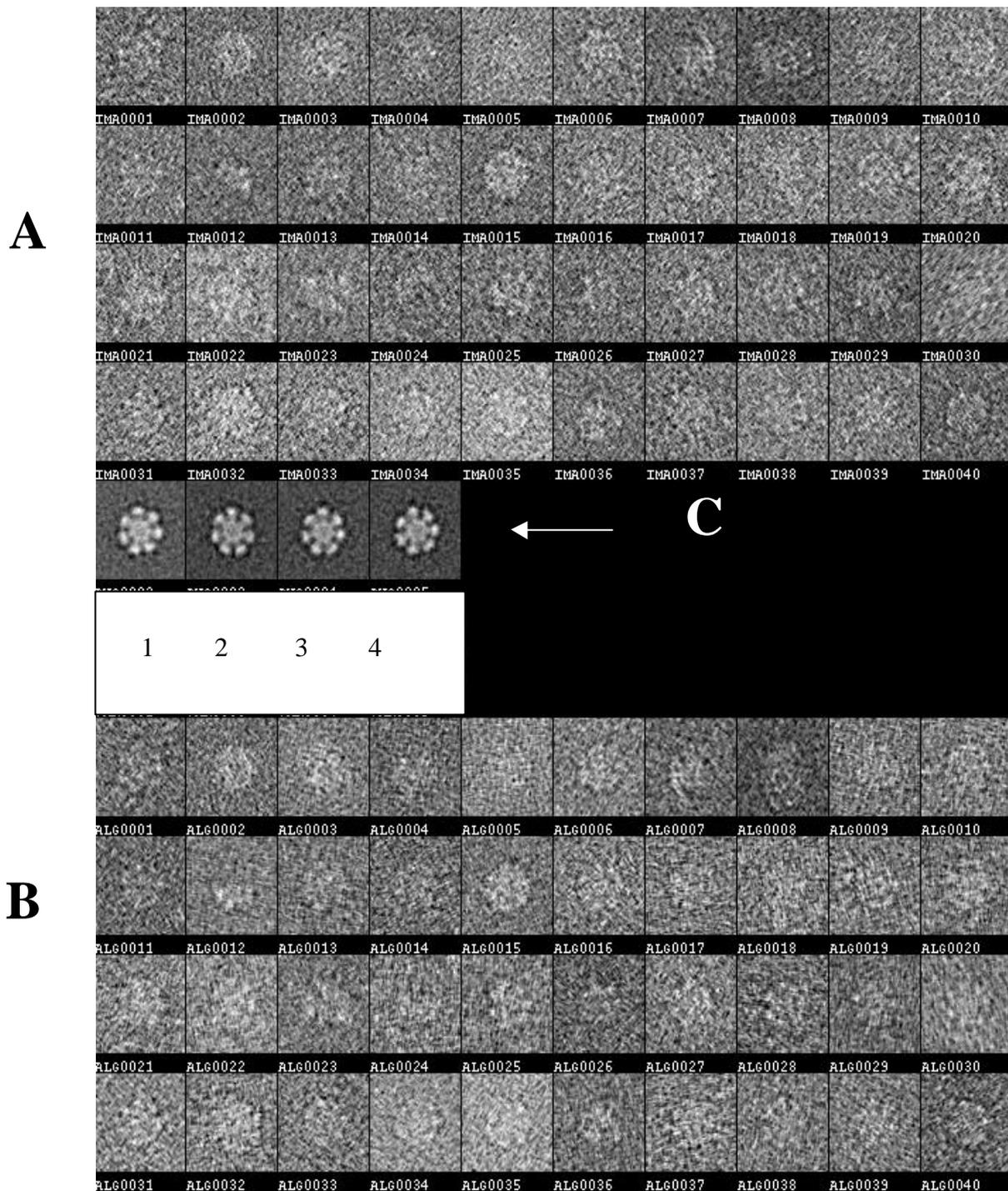


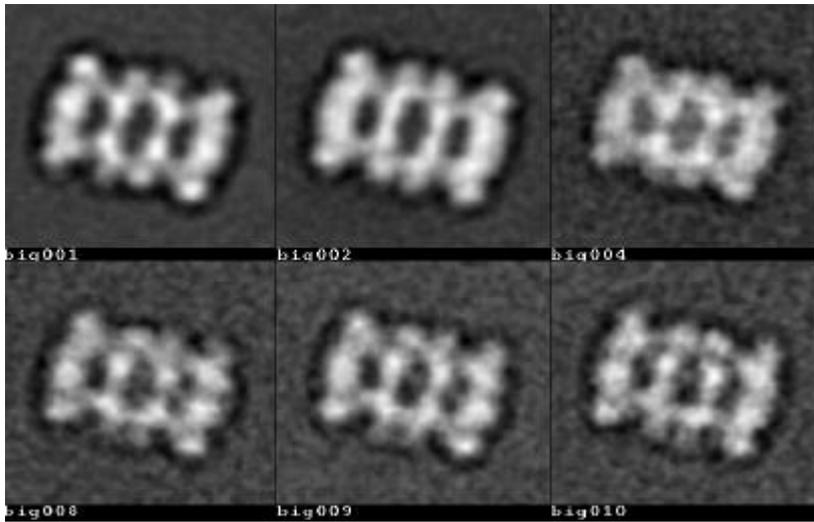
Fig. 5.4. Analyse bidimensionnelle des structures vues de dessus de l'anneau alpha. **A)** sélection des images appartenant à une même vue. **B)** réorientation des images sélectionnées par rapport à un même référentiel. **C)** image obtenue après superposition ou "moyennage" des images sélectionnées. Chaque image représente un cycle d'affinage supplémentaire. Image 1=1cycle, image 4=4 cycles.

Par analyse statistique des images digitalisées et réorientées, il est possible de classer chacune de ces images sur des critères de densités ou niveaux de gris. En effet, celles-ci possèdent les unes par rapport aux autres des informations nouvelles que d'autres n'ont pas. Le classement

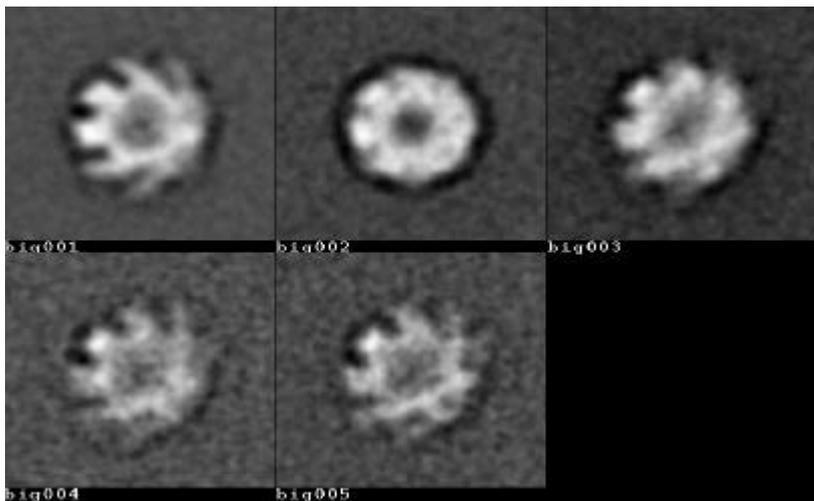
consiste à regrouper les images qui ont des informations très proches de manière à obtenir une image finale représentative de la classe. Celle-ci rassemble les informations données par l'ensemble des images. Cette classification ou analyse de correspondance s'effectue au moyen d'analyse statistique multivariée. Ces classes sont représentées à la figure 5.5. dont A) 6 pour

la vue de dessus et B) 5 pour les vues de côté du protéasome, et C) 5 pour l'anneau alpha libre. Ces classes sont très distinctes pour B et C.

A



B



C

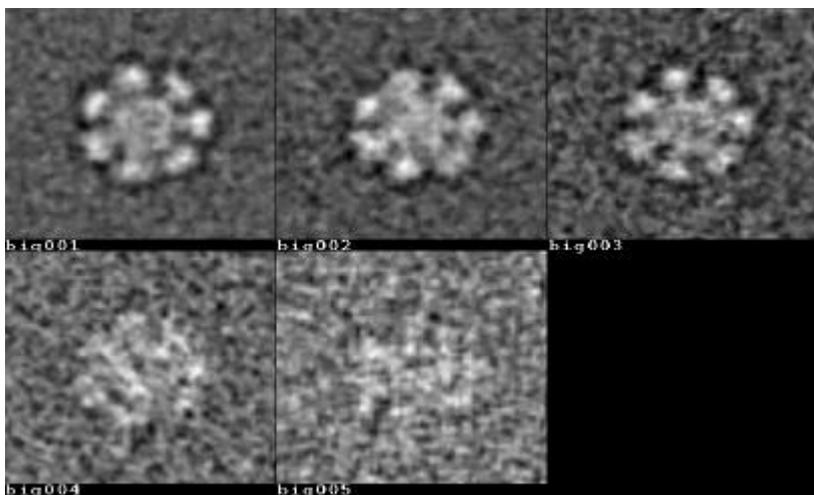


Fig. 5.5 Classification, par analyse de correspondance, des images bidimensionnelles qui représentent A) la vue de dessus B) la vue de côté C) le disque alpha du protéasome de mousse.

L'analyse statistique multivariée effectue enfin une classification hiérarchique ascendante. Elle rassemble et intègre les images des différentes classes dans une image finale bidimensionnelle de chacune des trois vues (fig. 5.6). Les figures 5.6 D, E et F présentent la carte des contours ou densité électronique des images bidimensionnelles. Elles visualisent le relief de chacune des structures.

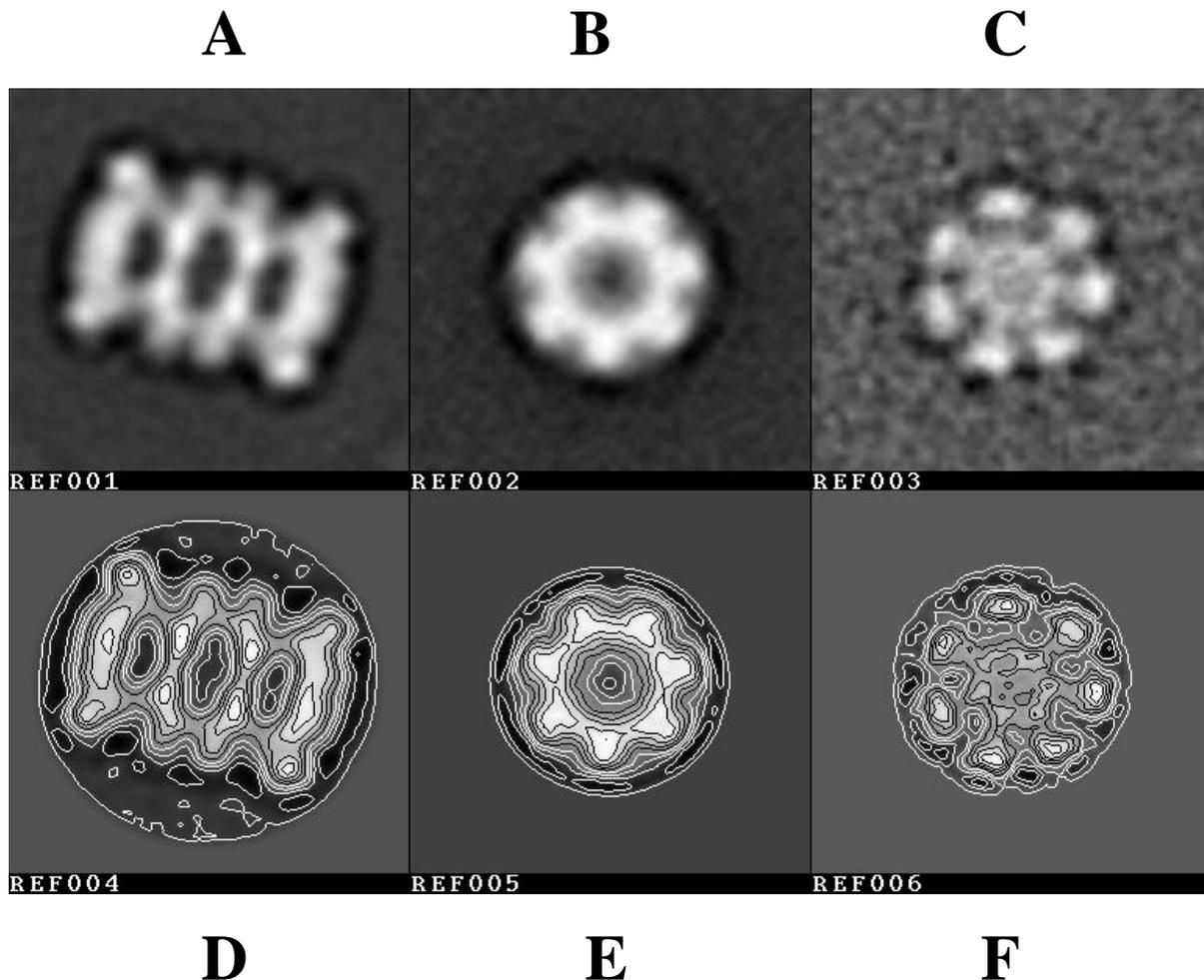


Fig. 5.6 Images bidimensionnelles finales des vues A) de dessus B) de côté C) de l'anneau alpha obtenues après classification hiérarchique ascendante par analyses statistiques multivariées. D), E) et F) représentent la carte des contours des vues 2-D correspondant respectivement aux images A), B) et C).

A ce stade, les images finales peuvent être comparées par superposition aux structures obtenues à partir du protéasome 20S cristallisé de levure. On constate (fig. 5.7 D, 5.8 D et 5.9 D) que la carte des contours des vues du protéasome de mousse (fig. 5.7.B, 5.8 B et 5.9 B) respecte fidèlement la position des sous-unités (fig. 5.7 C, 5.8 C et 5.9 C) dans la structure du protéasome de levure. Ceci est d'autant plus spectaculaire pour l'anneau alpha libre (fig. 5.9).

En dernier lieu, le volume de reconstruction est obtenu par analyse tridimensionnelle. Pour ce faire, chaque molécule observée en microscopie nécessite deux prises de vue à 45 degrés de différence d'angle (*tilting*). Cette opération augmente la quantité d'informations tridimensionnelles pour une même structure. Nous n'avons pas encore ces données, mais nous savons que cette protéine est symétrique, possède quatre anneaux à sept sous-unités, et que l'angle entre une sous-unité bêta et alpha est de 25 degrés selon les informations des protéasomes 20S cristallisés de la levure et de *T. acidophilum*.

Une fois les images bidimensionnelles et ces derniers paramètres introduits dans le programme de reconstruction volumique SPIDER, nous obtenons le volume de reconstruction par analyse tridimensionnelle présenté en figure 5.10. Les deux complexes 20S obtenus par *cryo-negative staining* et par cristallographie R-X sont identiques, si ce n'est le trou central qui est plus petit dans le cas du protéasome de mousse.

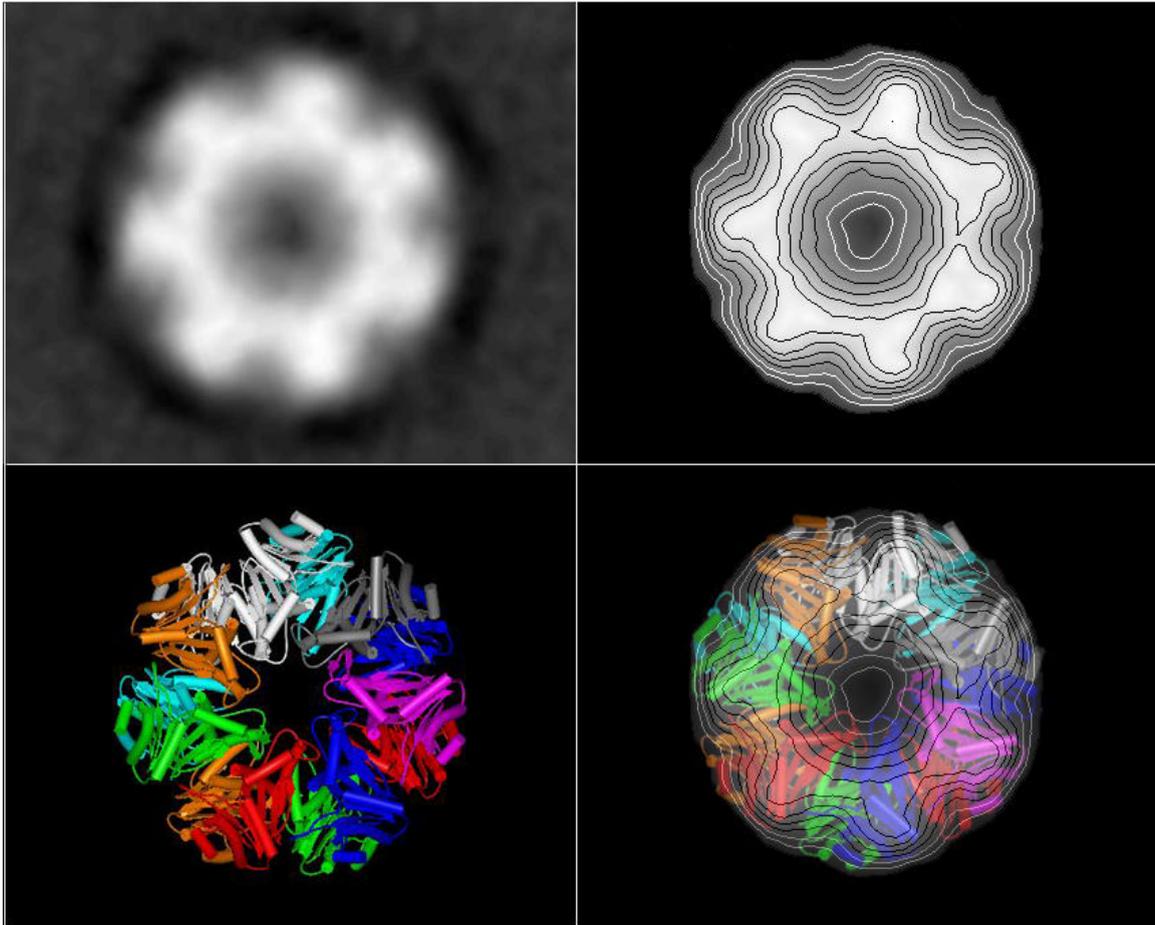


Fig. 5.7. A) Image bidimensionnelle, B) carte des contours de la vue de dessus du protéasome de mousse. C) vue de dessus de la modélisation de la structure cristalline du protéasome de levure. D) Superposition de B) et C).

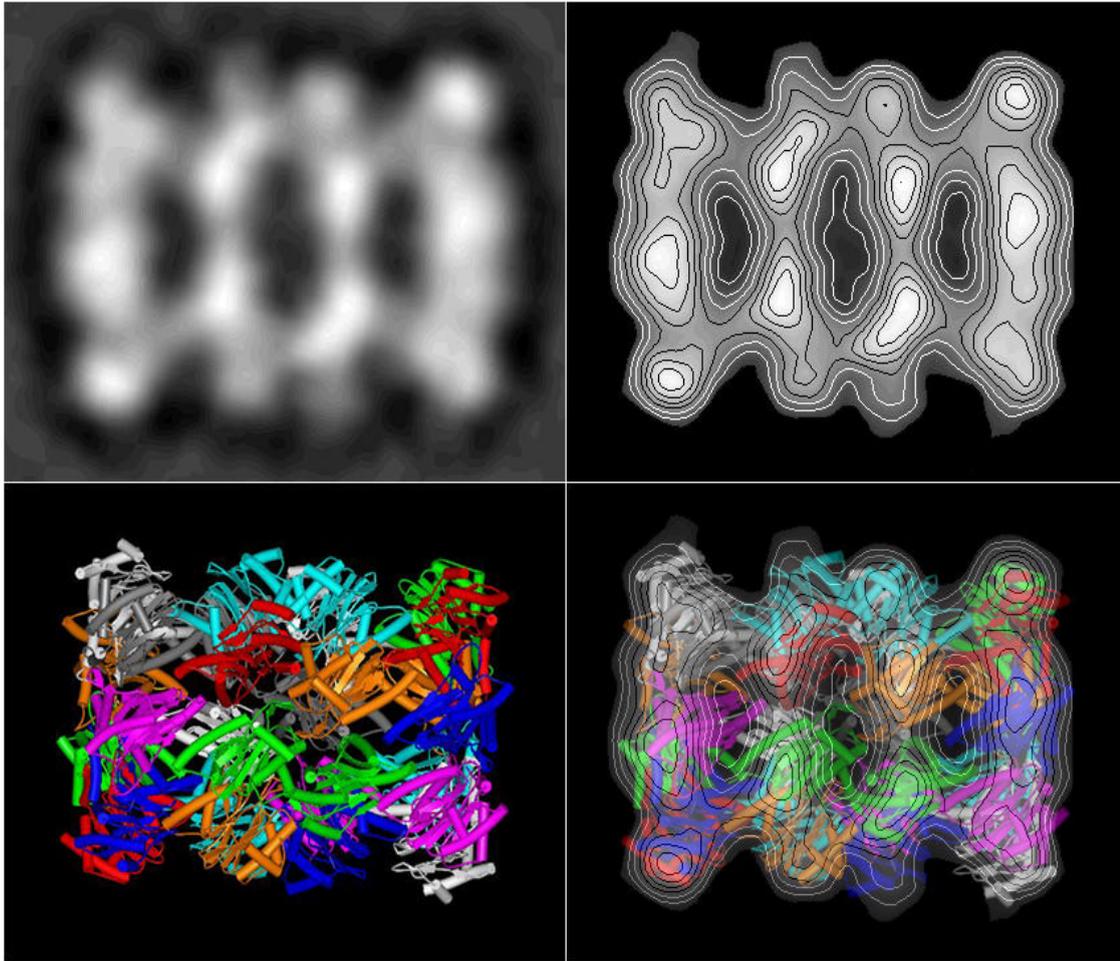


Fig. 5.8. A) Image bidimensionnelle, B) carte des contours de la vue de côté du protéasome de mousse. C) vue de côté de la modélisation de la structure cristalline du protéasome de levure. D) Superposition de B) et C).

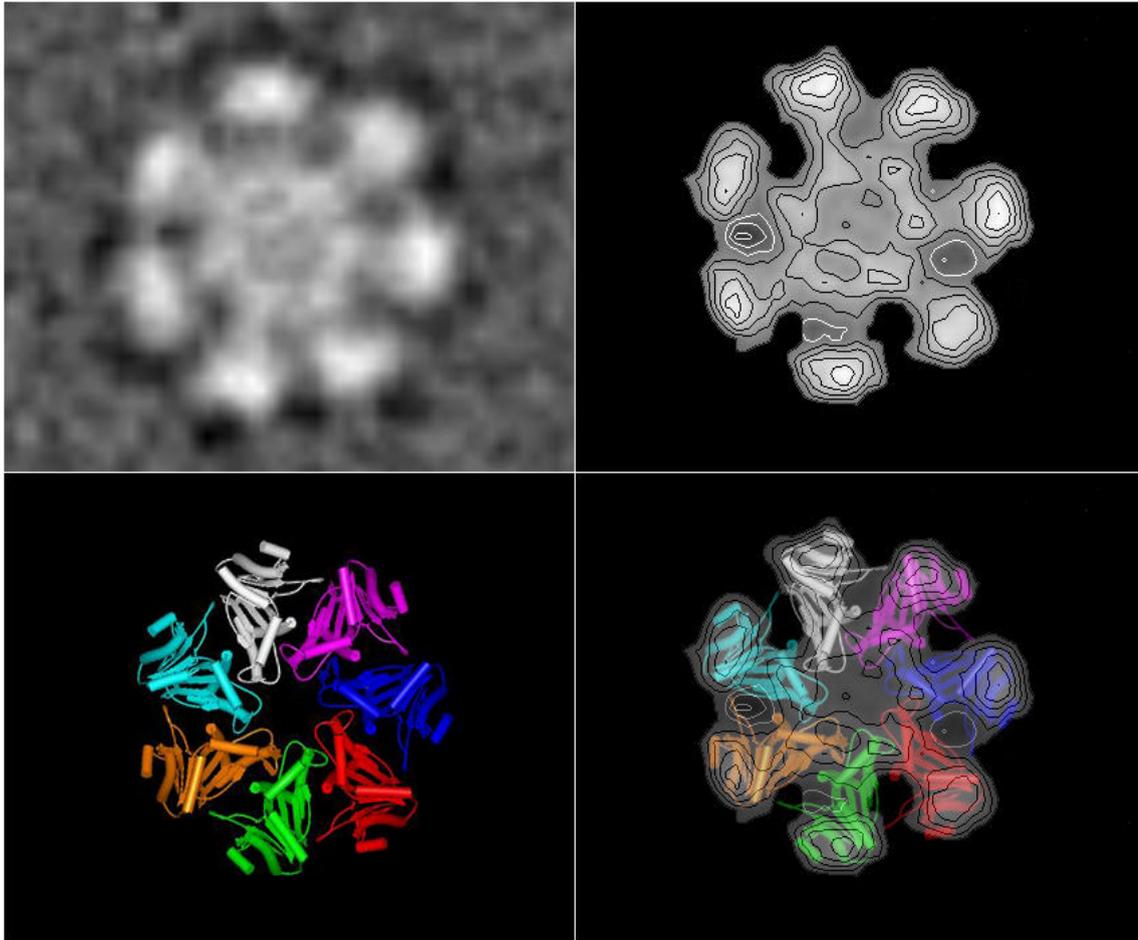


Fig. 5.9. A) Image bidimensionnelle, B) carte des contours de la vue du disque alpha du protéasome de mousse. C) vue de dessus de la modélisation de la structure cristalline du disque alpha du protéasome de la levure. D). Superposition de B) et C).

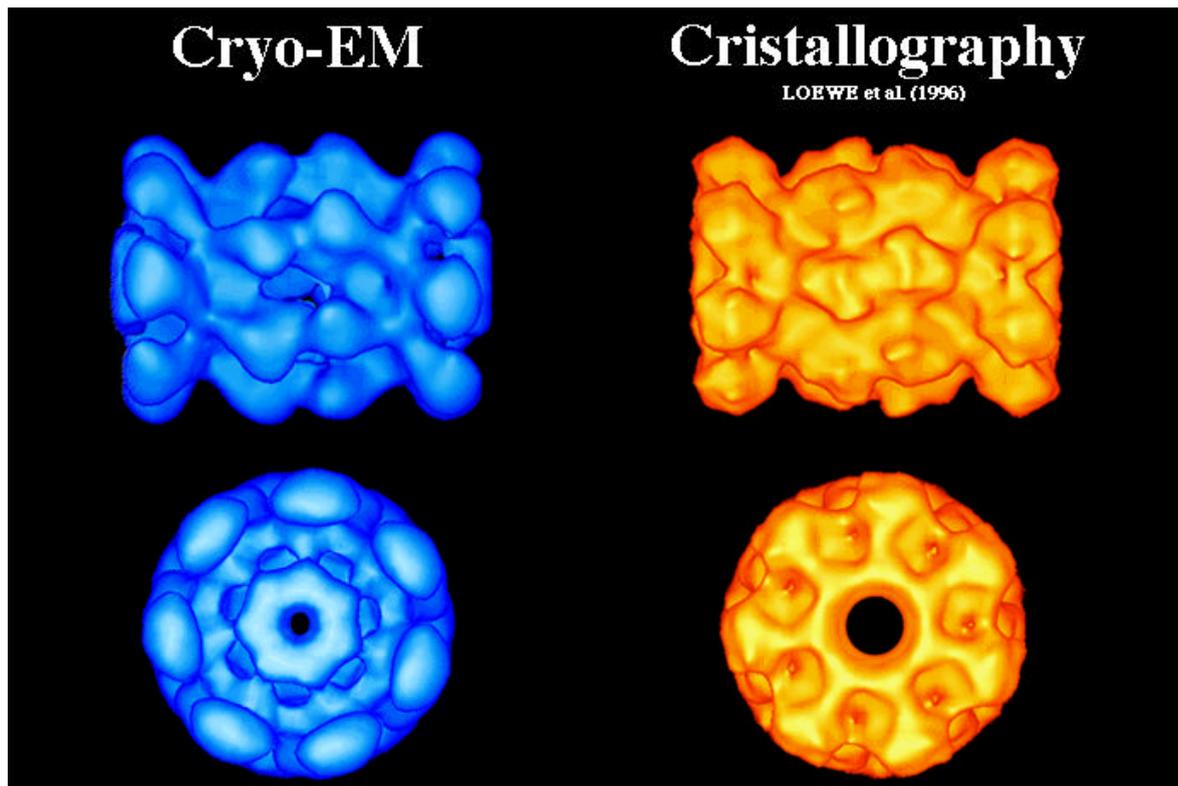


Fig. 5.10. Analyse en 3D. Image du volume de reconstruction du protéasome de mousse de *Physcomitrella patens* (en bleu) et du protéasome de levure obtenu par cristallographie (GROLL et al., 1997)

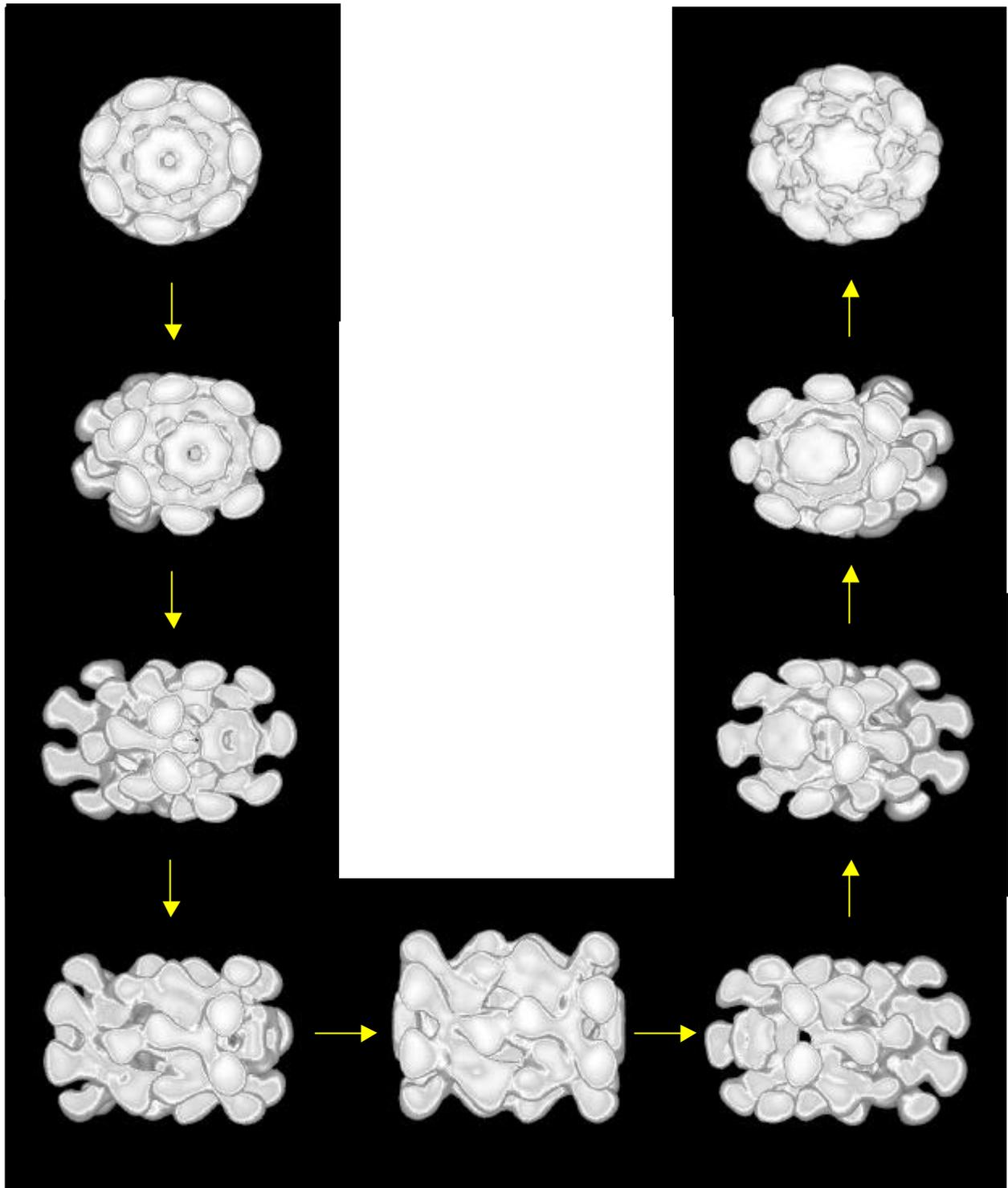


Fig. 5.11 Différentes vues du modèle de reconstruction 3-D du protéasome de *P. patens*. On constate que les deux anneaux alpha ne se trouvent pas dans le même état de conformation.

5.3 CONCLUSIONS

La technique de *cryo-negative staining* offre donc la possibilité de modéliser fidèlement une protéine et permet d'obtenir des informations spatiales de la molécule identiques à celles qu'obtient la cristallographie. Cette approche est très intéressante puisqu'elle nécessite moins de matériel biologique que la cristallisation. Elle est plus rapide et aussi fiable. Toutefois, sa résolution n'est pas aussi élevée que la diffraction R-X. On ne peut donc visualiser les structures secondaires des sous-unités du complexe.

La capacité d'extraction d'informations d'une image ou ce que l'on appelle le pouvoir de séparation en microscopie électronique est démontré de façon spectaculaire dans ce travail. Notamment, lorsque l'on arrive à modéliser l'anneau alpha à partir d'images qui semblent être inexploitable à notre œil (fig.5.4 A et B). De plus, on peut observer que le trou central du protéasome de mousse est plus petit. Si l'on recense les coordonnées cristallographiques de chacune des sous-unités alpha du protéasome de la levure, on constate qu'elles manquent pour les 3 à 7 acides aminés du côté de l'extrémité N-terminale. Ceci expliquerait un trou central plus grand pour la levure puisque le programme de modélisation n'a pas intégré ces données. Le fait que la diffraction R-X n'ait pas pu établir des coordonnées cristallographiques complètes pour certaines extrémités N-terminales, indique que celles-ci sont probablement instables. Cette instabilité doit également être présente *in vivo*.

L'information la plus remarquable concerne la polarité de la structure 20S du protéasome. En effet, la modélisation en trois dimensions du protéasome de *P. patens* nous a permis de visionner, en quelque sorte, l'état de conformation des deux anneaux alpha situés aux extrémités de la structure (fig.5.11). L'un se trouve dans un état plus relaxé que l'autre. Ceci laisse à penser que les substrats destinés à être dégradés ne peuvent accéder que par l'une des extrémités et que leurs produits de dégradation ne peuvent s'échapper que par l'autre extrémité. Or, le modèle du protéasome 26S tel que la plupart des chercheurs le conçoivent aujourd'hui offre aux substrats potentiels la possibilité d'accéder par les deux extrémités de la structure. Dans ce cas, la polarité fonctionnelle du protéasome 20S n'est plus respectée. Le modèle du protéasome 26S n'est donc pas cohérent.