

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la recherche Scientifique
المدرسة الوطنية العليا للعلوم الفلاحية – الحراش – الجزائر
Ecole Nationale Supérieure Agronomique - El Harrach - Alger
Ecole Doctorale: Biotechnologies Végétales مدرسة الدكتوراه: علم التقانة الحيوية النباتية



مذكرة
مقدمة لنيل شهادة الماجستير
في علم التقانة الحيوية النباتية
Biotechnologies végétales

من تقديم الطالب

خراز كريم

موضوع المذكرة

نمذجة بالتماثل ودراسة لبنية ووظيفة إنزيم Ferredoxin-nitrite reductase عند
نبات *Arabidopsis thaliana*
La modélisation par homologie et l'étude de la structure et la
fonction de Ferredoxin-nitrite reductase chez *Arabidopsis thaliana*.

أعضاء لجنة المناقشة:

رئيسا	المدرسة الوطنية العليا للعلوم الفلاحية-الحراش. الجزائر	أستاذ	خليفي خضر
مشرفا	المدرسة العليا للأساتذة - القبة. الجزائر	أستاذ	كاملي عبد الكريم
ممتحنا	المدرسة العليا للأساتذة - القبة. الجزائر	أستاذ محاضر	بوجنيبة مسعود
ممتحنا	المدرسة العليا للأساتذة - القبة. الجزائر	أستاذ محاضر	نور الدين رضوان

نوقشت يوم.....

ملخص

تم في هذه الدراسة بناء نموذج نظري ثلاثي الأبعاد لإنزيم Ferredoxin-nitrite reductase لنبات *Arabidopsis thaliana* ودراسة آلية عمله باستعمال طرق وبرامج البيومعلوماتية Bioinformatic tools، حيث اعتمدنا على طريقة النمذجة بالتماثل Homology modeling في بناء النموذج وذلك عبر أربع خطوات رئيسية:

الحصول على التابع الأميني للبروتين الهدف Target من قواعد البيانات الخاصة بذلك ثم البحث عن بنيات البروتينات القالب Templates الموافقة له. حيث تم اختيار البروتين 2akj المحددة بنيته تجريبيا كقالب والذي يملك نسبة التشابه الأعلى مع البروتين الهدف. بعد ذلك قمنا بمطابقة تتابع البروتين الهدف بنيته البروتين القالب 2akj التي تم اختيارها. ثم بناء 10 نماذج ثلاثية الأبعاد نظرية للبروتين الهدف بواسطة برنامج Modeller. وأخيرا تقييم دقة وصحة النماذج التي تم بناؤها، حيث تم اختيار النموذج الأكثر دقة و الذي يملك خصائص جيدة تقربه بشكل كبير جدا بنتائج البنيات المتحصل عليها تجريبيا (نسبة تشابه تقدر بـ 81% بين البروتين الهدف والقالب، 0.2 \AA قيمة RMSD، 99.5% من الأحماض الأمينية ذات توضع فراغي يقع في المنطقة المنطقية المسموح بها في مخطط Ramachandran plot، وقيمة Z-scores تساوي -10.37).

في الأخير قمنا بدراسة تحليلية لبنيته الثلاثية الأبعاد وموقعه الفعال ومحاوله فهم ودراسة آلية عمله.

الكلمات المفتاحية:

modeling, Modeller Homology Ferredoxin-nitrite reductase, *Arabidopsis thaliana*, program.

Résumé:

Dans cette étude a été fait une construction d'un modèle théorique tridimensionnel pour l'enzyme Ferredoxin-nitrite reductase chez *Arabidopsis thaliana*. l'étude de mécanisme de sa fonction en utilisant des programmes bioinformatiques.

On s'est basé sur la méthode la modélisation par homologie pour construire ce model, et cela par quater étapes essentielles.

L'obtention du séquençage du protéine but de Data base, et puis la recherche des moules des protéines qui lui correspond. La sélection du protéine 2akj identifié expérimentalement comme moule. On a fait un compatibilité entre le séquençage du protéine but et le moule 2akj qui a été sélectionné. Puis une construction de 10 modèles théoriques du protéine but par le programme modeller.

En fin, l'évaluation des modèles obtenus par différent paramètres ou le meilleur model a été choisi et qui possède de bonnes caractéristiques qui l'approche des moules expérimentales (le pourcentage de similarité 81%. RMSD= 0,2Å. 99,5 % d'acides aminés ont une conformation dans la région favorable dans la Ramachandran plot. Z-score = -10.37).

En fin on a procédé a une étude d'analyses pour le moule tridimensionnel et son site active et l'essaie de compréhension et l'étude de mécanisme de sa fonction.

Les mots clés :

Ferredoxin-nitrite reductase, *Arabidopsis thaliana*, Homology modelling, Modeller program.

Summary:

In this study, a theoretical three dimensional model of Ferredoxin-nitrite reductase from *Arabidopsis thaliana* was built its mechanism function analyzed using bioinformatics tools.

First, the sequence of target protein was obtained from relevant database, and then the template protein with known 3D structure and related sequence was identified (2akj).

Second, the target and template sequences were aligned. Third, the theoretical three dimensional models of target protein were built using modeller software. Fourth, the models obtained were evaluated using different criteria. The best model with high percentage similarity with the experimentally obtained structure (template) was chosen (81% amino acid sequence similarity, 0.2 Å as RMSD and -1.037 as Z-scores. The Ramachandran plot analysis showed that conformations for 99.5% of amino acid residues were within the most favoured regions.

Finally, the structure and mechanism function of best model were studied.

Keywords:

Ferredoxin-nitrite reductase, *Arabidopsis thaliana*, Homology modelling, Modeller program.

إهداء

إلى من كانا سببا في وجودي وتعليمي أمي
وأبي
إلى إخوتي وأقاربي
إلى مصابيح الدجى أساتذتي
اهدي ذا العمل

ك.خراز

شكر وتقدير

(...رب أوزعني أن أشكر نعمتك التي أنعمت علي وعلى والدي وأن أعمل صالحا ترضاه وأدخلني برحمتك في عبادك الصالحين...) الآية 19 سورة النمل

الحمد لله الذي وفقني في إتمام هذا العمل، وعلمني ما لم أكن أعلم وكان فضله عليّ عظيماً.

بصدق الوفاء والإخلاص أتشرف بتوجيه خالص الشكر وعظيم التقدير والامتنان والعرفان بالجميل إلى أستاذي الفاضل الأستاذ **كاملي عبد الكريم** على تكرمه بالإشراف على رسالتي، ولما لمستته من صدر رحب وتوجيه سديد ونصائح قيمة في إنجاز هذا العمل، أسأل الله العلي القدير أن يثيبه خير الثواب إنه سميع مجيب الدعاء.

كما أتوجه بخالص الشكر وعظيم التقدير والامتنان والعرفان بالجميل إلى الأساتذة الأفاضل: الأستاذ **خليفة لخضر** الذي تفضل بترؤس لجنة المناقشة، الأستاذ **بوجينية مسعود** والأستاذ **نو الدين** □ **ضوان** لقبولهما مناقشة المذكرة.

كما لا يفوتني أن أشكر صاحب الفضل الكبير **خراز خالد** وكل الأصدقاء والزملاء وكل من قدم لي يد المساعدة ولو بكلمة طيبة. واطمأن بالذکر زملائي في العمل وطلبة العلوم الطبيعية والحياة دفعة **2005**، طلبة مدرسة الدكتوراه علم التقانة الحيوية النباتية **2007**.

ك. خراز

1..... مقدمة عامة

الفصل الأول: الأساس النظري

6 1_ البيومعلوماتية **Bioinformatics**

6 1.1_ لمحة عامة عن البيومعلوماتية

7 2.1_ تطبيقات البيومعلوماتية

8 3.1_ مصادر معطيات الجزيئات البيولوجية

9 4.1_ بنوك معطيات البروتينات

9	1.4.1_ بنك معطيات البروتينات (PDB) Protein Data Bank
12	2.41_ بنك المعطيات NCBI
12	5.1_ الطرائق التجريبية المستعملة في تحديد بنية البروتين
13	1.5.1- الرنين المغناطيسي النووي NMR
14	2.5.1_ الدراسة البلورية بواسطة حيود الاشعة السينية X ray
16	6.1_ ملفات بنية البروتينات
16	7.1- البرامج المستعملة في عرض ملفات البروتينات
17	2_ البروتينات
17	1.2- العلاقة بين بنية ووظيفة البروتينات
19	3- التنبؤ ببنية البروتين Protein Structure Prédiction
21	1.3- النمذجة بالتماثل Homology modeling
24	3. 2_ تطبيقات النمذجة بالتماثل
25	4_ الخطوات الرئيسية للنمذجة بالتماثل
27	5_ تمثيل واهمية النيتروجين في النبات
27	1.5_ المرحلة الأولى
28	2.5_ المرحلة الثانية
30	6_ انزيم ال- Nitrite reductase
30	1.6_ مقارنة بين انواع ال- Nitrite-reductase
31	2.6_ نبذة تاريخية عن استخلاص ال- Nitrite-reductase
33	3.6_ مقارنة بين إنزيم ال- Nitrite reductase و Sulfite reductase
35	4,6_ بناء مجموعة ال- Siroheme
35	5,6_ موقع ارتباط مادة التفاعل النترت NO ₂ ⁻
36	6,6_ آلية انتقال الالكترونات المرجعة لمادة التفاعل النترت
37	7_ الفيرودوكسين Ferredoxin
38	8_ آلية ارتباط ال- Ferredoxin بال- Nitrite reductase

الفصل الثاني: الطرق و الوسائل

41	1_ النمذجة بالتماثل
----	---------------------

41.....	1.1_ جهاز الاعلام الالي.....
41.....	2.1_ بنوك المعطيات.....
41.....	1.2.1_ بنك معطيات البروتينات PDB.....
41.....	2.2.1_ بنك NCBI.....
42	3.1_ البرامج المستعملة.....
42.....	1.3.1_ برنامج Modeller.....
43.....	2.3.1_ برنامج Python.....
45	3.3.1_ برنامج Pymol.....
46	4.3.1_ برنامج VMD.....
47.....	2_ مراحل عملية النمذجة بالتمائل.....
47.....	1.2_ الخطوة الأولى: الحصول على تتابع البروتين الهدف وتحديد القالب المناسب.....
51.....	2.2_ الخطوة الثانية: مطابقة تتابع البروتين الهدف بالبروتين القالب.....
52.....	3.2_ الخطوة الثالثة: بناء النموذج.....
55.....	4.2_ الخطوة الرابعة: تقييم النموذج المتحصل عليه.....
55.....	1.4.2_ التقييم الداخلي Internal evaluation.....
55.....	1.1.4.2_ برنامج PROCHEK.....
55.....	2.4.2_ التقييم الخارجي Evaluation external.....
56.....	1.2.4.2_ برنامج PROSAll.....
57.....	2.2.4.2_ قيمة النتيجة Z-Score.....
58.....	3.2.4.2_ الجذر التربيعي لمتوسط المربعات RMSD Score.....

الفصل الثالث: النتائج و المناقشة

60.....	1_ النمذجة بالتمائل.....
60.....	1.1_ القالب او القوالب المناسبة.....
60.....	1.1.1_ أولاً: التتابع الأميني للبروتين الهدف.....
61.....	2.1.1_ ثانياً: تتابعات البروتينات المحددة البنية الشبيهة بالبروتين الهدف.....
64.....	2.1_ مطابقة تتابع البروتين الهدف بالبروتين القالب.....
69.....	3.1_ بناء النموذج.....
70.....	4.1_ تقييم النموذج المتحصل عليه.....

70.....	1.4.1_ حساب قيمة DOBE
72.....	2.4.1_ نتيجة المطابقة باستعمل RMSD
74.....	3.4.1_ تقييم البروتين النموذج المتحصل عليه بواسطة برنامج PROSAII
76....	4.4.1_ نتائج مخطط Ramachandran plot للبروتين النموذج المتحصل عليه
78.....	2_ الدراسة التحليلية لبنية Nitrite reductase
78.....	1.2_ البنية الأولية لل Nitrite reductase
79.....	2.2_ البنية الثانوية لل Nitrite reductase
81.....	3.2_ دراسة البنية الثالثية لل Nitrite reductase بواسطة برنامج Pymol
83.....	1.3.2_ المجال الاول
84.....	2.3.2_ المجال الثاني
84.....	3.3.2_ المجال الثالث
86.....	3_ الجزء غير البروتيني لل Nitrite reductase
87.....	1.3_ مجموعة ال [Fe ₄ S ₄]
88.....	2.3_ مجموعة Siroheme
90.....	3.3_ مادة التفاعل النترت
91.....	4_ الية عمل انزيم ال Nitrite reductase
95.....	الخلاصة
96.....	آفاق وتطلعات
98.....	المراجع
113.....	الملحق

الاختصاات

CATH : Class Architecture Topology Homology

cDNA: complementary Deoxyribonucleic acid

DNA: Deoxyribonucleic acid

DOPE: Discrete Optimized Protein Energy

3D: three dimensional

EPR: Electron paramagnetic resonance

FAD: Flavin Adenine Dinucleotide

FNR: Ferredoxin: NADP oxidoreductase

FTR: Ferredoxin:thioredoxin reductase

Fd: Ferredoxin
GOGAT: Glutamate synthase
Kd: kilo Dalton
MmCIF: Macromolecular Crystallographic Information File
NCBI: National Center for Biotechnology Information
NIR: Ferredoxin:nitrite oxidoreductase
NR: Nitrate reductase
NMR: Nuclear Magnetic Resonance
PDB: Protein Data Bank
PIR: Protein Information Resource
PSI : Photosystem I
PSVS: Protein Structure Validation Software
RCSB: Research Collaboratory for Structural Bioinformatics
RMSD: Root Mean Square Deviation
RNA: Ribonucleic acid
SCOP : Structural Classification of Proteins
SiR: Ferredoxin:sulfite reductase
Uniprot : Universal Protein Resource
VMD: Visual Molecular Dynamics
SCOP : Structural Classification Of Proteins



مقدمة عامة:

يأتي النتروجين في المرتبة الرابعة بعد الكربون، الهيدروجين، والأكسجين في نسبة تركيب الكائن الحي، إذ يدخل في تركيب مركبات هامة مثل البروتينات، الأحماض الأمينية، الأحماض النووية، وبعض منظمات النمو وفي الكثير من الفيتامينات، ونظرا لهذه الأهمية الكبيرة فللنتروجين دور ودخل كبير في جل التفاعلات الكيميائية الحيوية التي تشكل عصب الحياة، كما يعتبر في غالب الأحيان العنصر المحدد الذي يعيق نمو النبات (Fowden ., 1979). ورغم توفر النتروجين بكميات كبيرة في الهواء تصل إلى 79 % إلا أن النبات لا يستفيد منه مباشرة، لذا يتوجب إضافته للتربة باستمرار على شكل أسمدة كيميائية أو عضوية، أين يدخل النتروجين في تكوين المركبات الهامة في النبات كالأحماض الأمينية والبروتينات والكلوروفيل وفي تكوين الأنسجة حديثة النمو في الأفرع والأوراق والجذور والثمار والنبور، لذا يعتبر من العناصر الهامة في نمو المحاصيل الزراعية.

تمتص غالبية النباتات الراقية النتروجين من التربة في صورة نترات NO_3^- (Hirsch & Sussman., 1999; Crawford & Glass ., 1998) ولكن النتروجين في صورته هذه لا يستغل مباشرة من قبل النبات إلا بعد أن يختزل إلى أمونيا قبل أن يدخل في تركيب المركبات النتروجينية الموجودة في النبات فيما بعد، حيث أن عملية اختزال النترات NO_3^- إلى الأمونيا NH_4^+ تتم في مرحلتين:

تتم المرحلة الأولى داخل السيتوبلازم، حيث بعد دخول النترات NO_3^- إلى داخل الخلايا يتم اختزالها إلى نترت (NO_2^- إلى NO_3^-) بواسطة إنزيم الـ Nitrate reductase.

أما المرحلة الثانية فهي تأتي بعد تحويل النترات إلى نترت (تعتبر سامة بكميات معينة)، تهاجر هذه الأخيرة إلى داخل الصانعات الخضراء في الأنسجة اليخضورية وإلى داخل البروبلاست في الأنسجة غير اليخضورية، أين تُختزل إلى أمونيا NH_4^+ بواسطة إنزيم (Rideout et al., 1994) Ferredoxin-nitrite reductase، تدمج الأمونيا بعد ذلك في بناء الأحماض الأمينية بواسطة إنزيمين آخرين هما Glutamine synthetase و Glutamate synthase (Vega et al., 1991).

تم تحديد وعزل الجينات المشفرة لإنزيم Ferredoxin-nitrite reductase للعديد من النباتات الراقية وتم استنساخها وتحديد تتابعها، حيث تم عزل الجينات المشفرة له مثلا عند نبات السبانخ

استخلاص إنزيم Ferredoxin-nitrite reductase من عدة أنواع نباتية مثل السبانخ (Lahners et al., 1988) Maize و الذرى (Back et al., 1988) Spinach، كما تم (Serra et al., 1982) Barley الشعير، (Vega et al., 1977) Spinach البازلاء، (Bowsher et al., 1988) Cyanobacteria مثل البكتيريا الزرقاء، (Merchan et al., 1995) *Phormidium Laminosum* و *Thermosynechococcus* مثل *Bacillus cereus* Bacteria والبكتيريا، (Nakamura et al., 2002) *Elongates* و (Ivanova et al., 2003) *Pirellula* sp و (Gloeckner et al., 2003) الطحالب، (Archibald et al., 2003) *Bigelowiella natans* مثل Algae.

دلت الدراسات السابقة على أن □ إنزيم Ferredoxin-nitrite reductase عند كل من النباتات، الطحالب، البكتيريا الزرقاء هو بروتين ذو سلسلة ببتيدية واحدة Monomeric وبوزن جزيئي يقارب الـ 63 كيلو دالتون (Back et al., 1988; Hilliard et al., 1991)، يحتوي على مجموعتين غير بروتينيتين Prosthetic groups في موقعه الفعال Active Site، وهما مجموعة واحدة من الـ Siroheme (Vega et al., 1977; Ida et al., 1986) و مجموعة واحدة من $[Fe_4-S_4]$ (Hirasawa et al., 1987a; Ip et al., 1990).

سمحت التقنيات التجريبية الحديثة المعتمدة في تحديد التتابعات الجينية بمسح كامل الجينوم للعديد من الكائنات الحية، من بينها الإنسان ونبات *Arabidopsis thaliana*، وهي الآن تسير بوتيرة متسارعة. وهذا ما أدى إلى زيادة مطردة في عدد التتابعات البروتينية المحددة، لكن هذه السرعة في تحديد الكم الهائل من التتابعات الجينية والأمنية لم تتبعها نفس الوتيرة في تحديد البنيات ثلاثية الأبعاد للبروتينات.

إن تحديد البنية الثلاثية الأبعاد هي المفتاح الأساس لفهم وظيفة البروتينات، والتي تعتبر من المسائل والتحديات الأساسية الحالية التي تواجهها البيولوجيا الجزيئية، أين استخدمت طريقتي الرنين النووي المغناطيسي و الأشعة X التجريبتين للتغلب على هذا المشكل، لكن هاتين الطريقتين رغم نجاعتهما وجودة نتائجهما، إلا أنهما لم يحلا مشكل تحديد البنية البروتينية كليا لعدة أسباب حسب (Jones et al., 2000):

1- طريقتي الرنين النووي المغناطيسي والأشعة X متوفرة في عدد قليل جدا من المخابر لتكلفتها الباهظة.

2- بعض البروتينات يصعب تحديد بنيتها البلورية بسهولة نظرا لصعوبة استخلاصها وتنقيتها بشكل كامل.

3- تحديد البنية البلورية لبروتين واحد تجريبيا بواسطة الأشعة X يستغرق وقتا طويلا من عدة شهور إلى عدة سنوات.

4- تحديد البنية ثلاثية الأبعاد للبروتينات بواسطة الرنين النووي المغناطيسي وإن كانت أسرع من الأشعة X إلا أنه لا يمكن تطبيقها حاليا إلا على البروتينات التي لها عدد أحماض أمينية أقل من 100 حمض أميني.

5- هناك فجوة كبيرة جدا بين العدد الكبير للتتابعات البروتينية المحددة والمقدرة (14-Apr-2009) بحوالي 7527796 تتابع أميني (www.ebi.ac.uk/uniprot)، والعدد القليل نسبيا من البنيات ثلاثية الأبعاد للبروتينات التي تم تحديدها تجريبيا والتي تقدر بحوالي 57133 بنية (<http://www.rcsb.org/pdb/static>).

6- كنتيجة لمشاريع تحديد تتابع الجينوم فإن التتابعات الأمينية للبروتينات التي يتم تحديدها تتزايد بشكل كبير جدا.

7- إن افترضنا أنه سيتم تحديد كل البنيات الثلاثية الأبعاد لكل البروتينات المحددة التتابع تجريبيا بواسطة الرنين النووي المغناطيسي والأشعة X فإنها تتطلب أكثر من 500 عام.

ونتيجة لهذه الأسباب والعوائق أصبح الإلتجاء إلى أدوات البيومعلوماتية Bioinformatics أمراً لا مفر منه. حيث طوّر العلماء برامج حاسوبية تعتمد أساسا على الخوارزميات الرياضية القادرة على التنبؤ بالبنية ثلاثية الأبعاد للبروتين والتي تعرف ببرامج التنبؤ بالبنية ثلاثية الأبعاد .Protein Structure Prediction

التنبؤ ببنية البروتين من أقدم المسائل التي واجهت البيومعلوماتية Bioinformatics المختصة بدراسة البنية، وقد استعملت عدة طرق مختلفة للتغلب على هذا المشكل على مدى السنين، وكانت أكثر هذه الطرق نجاحا تلك التي اعتمدت على المعلومات الخاصة بالعلاقة بين التتابع الأميني و البنية ثلاثية الأبعاد (Sali et al., 1995; Moult et al., 2005)، حيث تعمل هذه الطريقة على بناء نموذج نظري للبنية ثلاثية الأبعاد للبروتين انطلاقاً من المعطيات الأولية لتتابعه الأميني والتي تدعى بطريقة النمذجة بالتماثل Homology modeling.

على الرغم من تحديد الكثير من التتابعات الجينية والأمينية لإنزيم Ferredoxin-nitrite reductase للكثير من النباتات الراقية المتوفرة في قواعد المعطيات للجزيئات الحيوية كبنك NCBI. إلا أنه لم يتم تحديد إلا بنية واحدة فقط تجريبيا بواسطة الأشعة السينية لـ Ferredoxin-nitrite reductase لنبات السبانخ الموجودة في بنك المعطيات PDB والبنوك الأخرى تم تحديدها.

ومساهمة منا لمعرفة البنية ثلاثية الأبعاد وفهم آلية عمل هذا الإنزيم المهم جدا في تمثيل النترات عند نبات *Arabidopsis thaliana*. وباستعمال أدوات وتقنيات البيومعلوماتية tools Bioinformatics، واعتمادا على النمذجة بالتماثل Homology modeling كطريقة للتنبؤ بالبنية الثلاثية الأبعاد، قمنا ببناء 10 نماذج بنيوية لـ Ferredoxin-nitrite reductase التي تم تقييمها بعدة برامج بيومعلوماتية واختيار النموذج الأحسن بينها. بعد ذلك قمنا بدراسة تحليلية تفصيلية لمختلف مستويات بنية هذا الإنزيم ومحاولة فهم آلية عمله اعتمادا على الدراسات السابقة لهذا الإنزيم عند أنواع نباتية أخرى و الإنزيمات الشبيهة.

الأحاسيس

النظري

1- البيومعلوماتية Bioinformatics

1.1_ لمحة عامة عن البيومعلوماتية

شهدت العلوم البيولوجية تطورا كبيرا خلال العقدين الماضيين، مما أدى إلى إنتاج كميات ضخمة جدا من المعطيات البيولوجية (Reichhardt et al., 1999)، فعلى سبيل المثال نجد أن بنك التسابغات الجينية Genbank في عام 2009 يحتوي على حوالي 100 مليون تتابع نيكليوتيدي، وبنك بنيات البروتينات PDB يحتوي على أزيد من 51491 بنية ثلاثية الأبعاد محددة تجريبيا، وبنك المعطيات Uniprot يحتوي على 7.5 مليون تتابع بروتيني. هذه المعطيات تتطور بمقدار يقارب الضعف كل خمسة عشر شهرا (Benson et al., 2000)، ومن الواضح أنه إن لم يتم تخزين و تنظيم و تحليل هذا الكم الهائل من المعطيات فإنها لن تكون مفيدة ومتاحة بقدر كاف، لذا كان لزاما على العلماء والباحثين إيجاد تقنيات وبرامج تساعد على التغلب على هذا المشكل، هذا ما أدى إلى ظهور علم جديد اسمه البيومعلوماتية Bioinformatics والذي يعني استعمال وتطبيق أحدث التقنيات والبرامج الحاسوبية لدراسة، فهم، تحليل و تنظيم المعلومات و المعطيات والبيانات المتعلقة بالجزئيات البيولوجية وذلك بدمج كل من علوم البيولوجيا، الرياضيات، الفيزياء و الإحصاء وبشكل أساسي علوم الحاسوب مما جعل من البيومعلوماتية حقلا خصبا جلب اهتمام الكثير من الباحثين والدارسين في العقدين السابقين.

وفي الحقيقة فإن التوسع الكبير الذي شهده عالم البيومعلوماتية يرجع أساسا إلى التطور المتسارع لتكنولوجيا الحاسوب وبالخصوص البرامج الإلكترونية وآليات ووسائل التخزين الإلكترونية التي تتطور بشكل متسارع يوما بعد يوم، وكذا إلى شبكة الإنترنت التي تسمح للباحثين والدارسين بالحصول على المعطيات البيولوجية بشكل أسهل في شتى أنحاء العالم كما تمنحهم إمكانية إدخال المعطيات الجديدة عن الجزئيات الحيوية المتحصل عليها في مخابر بحوثهم إلى بنوك المعطيات البيولوجية كبنك معطيات بنيات البروتينات PDB (Bernstein et al., 1977;) (Berman et al., 2000).

وبما أن تحليل ودراسة هذه المعطيات البيولوجية هي أهم أهداف البيومعلوماتية، تم تطوير برامج حاسوبية تساعد على تحليل هذه المعطيات، فعلى سبيل المثال يمكن استنتاج بنية ووظيفة بروتين تم تحديده بتابعه الأميني حديثا وذلك بمقارنته بالبروتينات المحددة البنية تجريبيا، ومن بين أهم هذه البرامج نجد برنامج FASTA (Pearson et al., 1990) وكذا برنامج PSI-PLAST (Altschul et al., 1997) والذي تمكن من تحديد نقاط التشابه والاختلاف بين التسابغات الأمينية للبروتينات.

سابقا كانت الدراسات ونتائج التجارب البيولوجية المعمولة في المخبر تجرى بشكل معزول وعادة ما تقارن بعدد قليل من نتائج التجارب المخبرية ذات الصلة المباشرة، لكن بظهور البيومعلوماتية أصبح من الممكن دراسة وتحليل ومقارنة نتائج معطيات بيولوجية كثيرة مجتمعة كمقارنة تتابعات جينية أو أمينية أو بنيوية لعدد كبير جدا من البروتينات على الحاسوب وفي وقت قصير جدا.

2.1_ تطبيقات البيومعلوماتية

كما تبين سالفا فإن البيومعلوماتية ظهرت في الأعوام الأخيرة كفرع أساسي وفعال في البيولوجيا الحديثة عموما والأبحاث الطبية الجزيئية على وجه الخصوص، وذلك لاستعمالاتها المختلفة التي تلبي حاجات الباحثين والدارسين في هذا المجال.

إن الكم الهائل والمنظم من المعلومات التي توفرها البيومعلوماتية في قواعد بياناتها و المتعلقة بالجينات وتعبيراتها وكذا بنيات الجزيئات الحيوية مكن من إتاحة، تسهيل، فهم، وتحويل هذه المعلومات إلى الباحثين في هذا المجال، حيث يمكن للباحثين في المجال الطبي مثلا الاستفادة بشكل كبير من المعلومات الخاصة بالتتابعات الجينية وتعبيراتها المتوفرة في قواعد البيانات البيولوجية، وتساعد الباحثين في حقل الصيدلة من تصميم الأدوية Drug Design، كما تمكن البيولوجيين من الإجابة على العديد من التساؤلات البيوكيميائية والبيوفيزيائية المتعلقة بتشابه واختلاف الإنطواءات للبنيات البروتينية وربطها بوظائفها (Martin et al., 1998) والتي تؤدي إلى فهم مدى اختلاف وتشابه الطرق الأيضية عند الكائنات الحية (Tatusov et al., 1997; Tatusov et al., 1996).

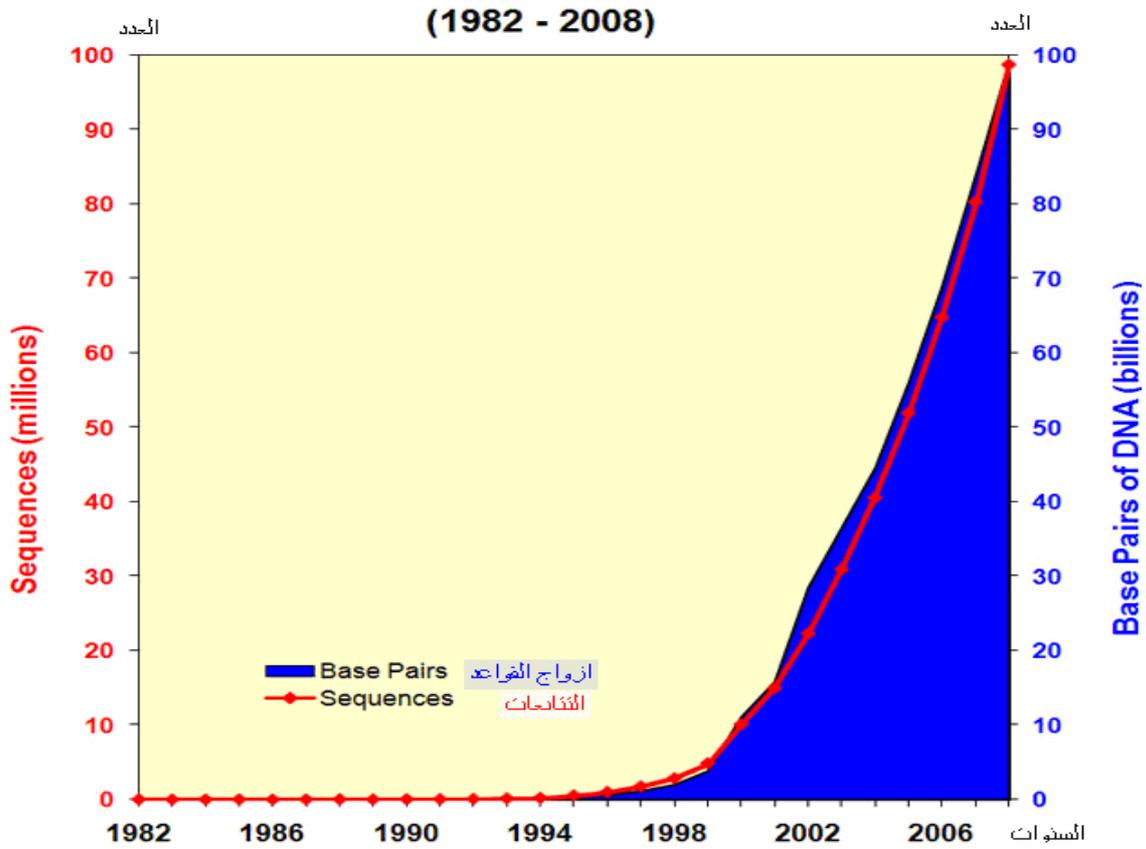
كما يمكن ربط المعلومات الخاصة بالتعبير الجيني ببنية البروتين ووظيفته في وضع تصنيف للبروتينات (Jansen et al., 2000) و توضعها في الخلية (Uetz et al., 2000. Ross-) (Macdonald et al., 1999).

3.1_ مصادر معطيات الجزيئات البيولوجية

تتضمن المعطيات والبيانات البيولوجية تتابعات وبنيات الحمض النووي الريبي منقوص الأكسجين DNA، تتابعات الحمض النووي الريبي RNA، تتابعات الأحماض الأمينية

للبروتينات وبنيات الجزيئات الحيوية الأخرى حيث يعتبر بنك الجينات Genbank مستودعا هاما للتتابعات الجينية والذي يحتوي على عدد كبير جدا وهو في تطور ونمو متسارع منذ نشأته عام 1982 (98,868,465 تتابع عام 2008) كما هو موضح في (الشكل 1). بحيث يمكن تنزيل أي تتابع جيني تم تحديده مجانيا من على موقع هذا البنك على شبكة الإنترنت <http://www.ncbi.nlm.nih.gov/Genbank/>

بالنسبة للتتابعات الأمينية للبروتينات فإنها متوفرة في عدد من البنوك الخاصة كبنك المعطيات NCBI، بنك SWISS-PROT (Bairoch et al., 1999) و بنك PIR (McGarvey et al., 2000).



الشكل (1): يوضح تطور عدد التتابعات النيكلوتيدية في بنك Gnbank من الموقع

<http://www.ncbi.nlm.nih.gov/Genbank/genbankstats.html>

أما البنيات ثلاثية الأبعاد المحددة تجريبيا للجزيئات الحيوية بصفة عامة وللبروتينات بصفة خاصة فهي ممثلة بعدة بنوك أهمها بنك المعطيات PDB (Berman et al., 2003).

4.1_ بنوك معطيات البروتينات

1.4.1_ بنك معطيات البروتينات (PDB) Protein Data Bank

بنك معطيات البروتينات PDB مصدر جد مهم للجزيئات الحيوية، أنشأ عام 1971 بالمخبر الوطني لكروهافن من طرف كل من Walter Hamilton و Edgar Meyer، وهو عبارة عن أرشيف ضخم للبنيات ثلاثية الأبعاد الجزيئات الحيوية الكبيرة كالبروتينات والأحماض النووية، حيث يعتمد على الاستغلال الفعال لآخر التطورات العلمية وكذا الوسائل والتقنيات المعتمدة في هذا الميدان لتحديد البنية تجريبيا، كالرنين النووي المغناطيسي [MRN](#) وحيود الأشعة السينية-[X-Ray](#)، و يشارك في تحديد هذه البنيات باحثون من اختصاصات مختلفة يعملون بشكل متكامل (بيولوجيون، فيزيائيون وكيميائيون)، كما يعتمد بنك الPDB على تكنولوجيا الحاسوب أساسا في معالجة و تخزين واسترجاع بيانات ومعطيات هذه الجزيئات الحيوية (Berman et al., 2003) ، حيث يقدم المعلومات والتحليل الخاصة بكل بنية على ثلاثة أقسام (Berman et al.,) 2006:

1- معلومات حول المحتوى الكيميائي ككتابع وعدد الأحماض الأمينية و المجموعات غير البروتينية المرتبطة بالبروتين وكذا الروابط الكيميائية.

2- معلومات خاصة بنمط التجربة المستعملة في تحديد البنية ثلاثية الأبعاد (الرنين النووي المغناطيسي أو الأشعة X) مع تحديد دقة التصميم (Resolution).

3- بعض الأوصاف النوعية الخاصة بالبنية (الأولية، الثانوية، الثالثية و الرابعة).

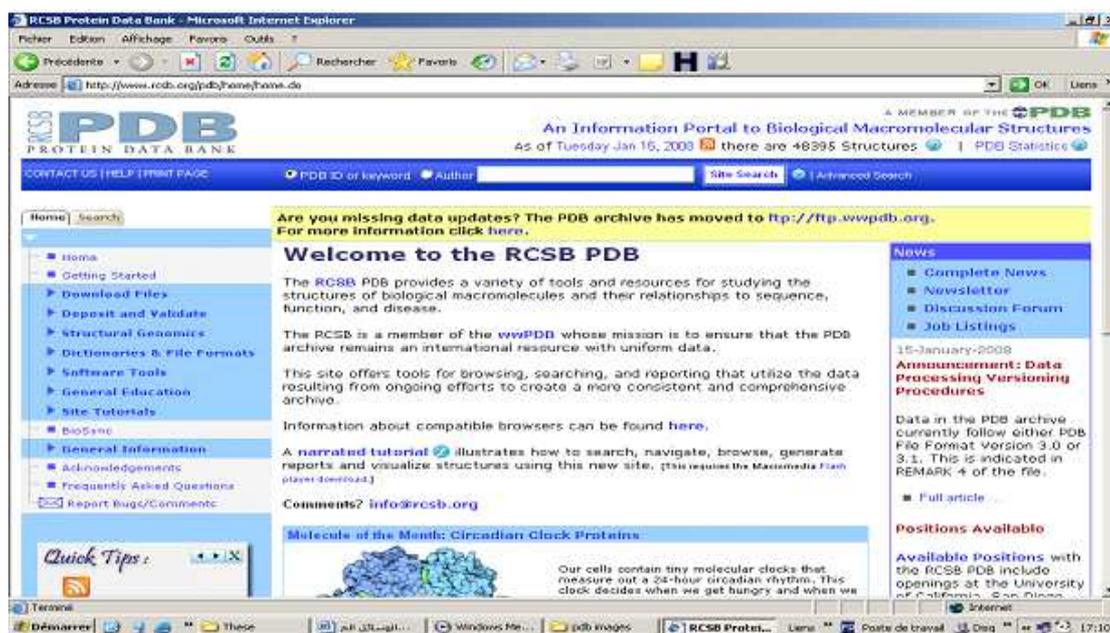
ومنذ تاريخ إنشاء هذا البنك تطور أرشيفه بصورة متسارعة جدا، حيث انتقل من 7 بنيات تم إدخالها عام 1973، إلى أكثر من 50000 وحدة بنوية عام 2009، كما هو موضح في (الجدول 1) و(الأشكال 2،3،4). هذا النمو في حجم الأرشيف تمثل في الزيادة في كل من محتوى المعطيات وكذا التعقيد الذي ساير الجزيئات المدخلة.

الجدول (1): يوضح بنيات الجزيئات الحيوية المخزنة في بنك PDB المحددة بالرنين النووي والأشعة X حتى 2009.

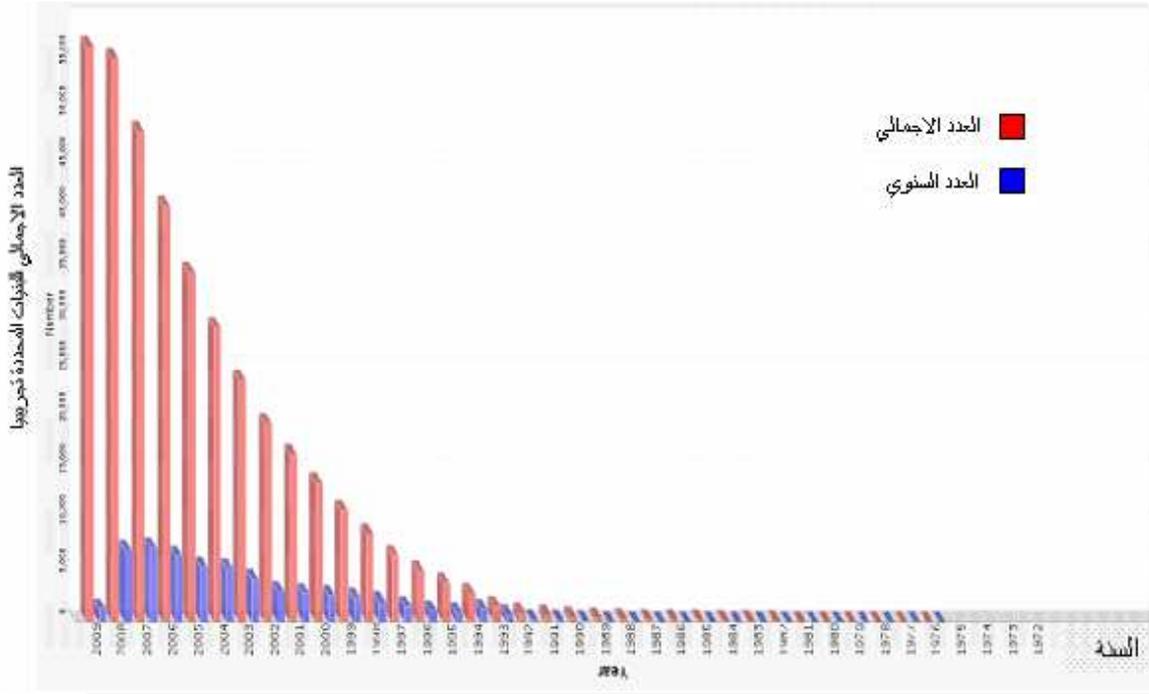
المجموع	جزيئات أخرى	معقد Protein/NA	الأحماض النوية	البروتينات	الجزيئات الحيوية التجربة
51626	17	2216	1168	48225	الأشعة السينية
8018	6	150	869	6993	الرنين النووي المغناطيسي

إن التطور الكبير لـ PDB أدى إلى توسيع نطاق مستعملي البنك، حيث انتقل البنك من مجرد ممول بسيط للبيانات الهيكلية للجزيئات إلى مصدر رئيسي للبيانات وكل المعلومات البيولوجية والكيميائية الخاصة بالجزيئات الحيوية (Westbrook et al., 2005).

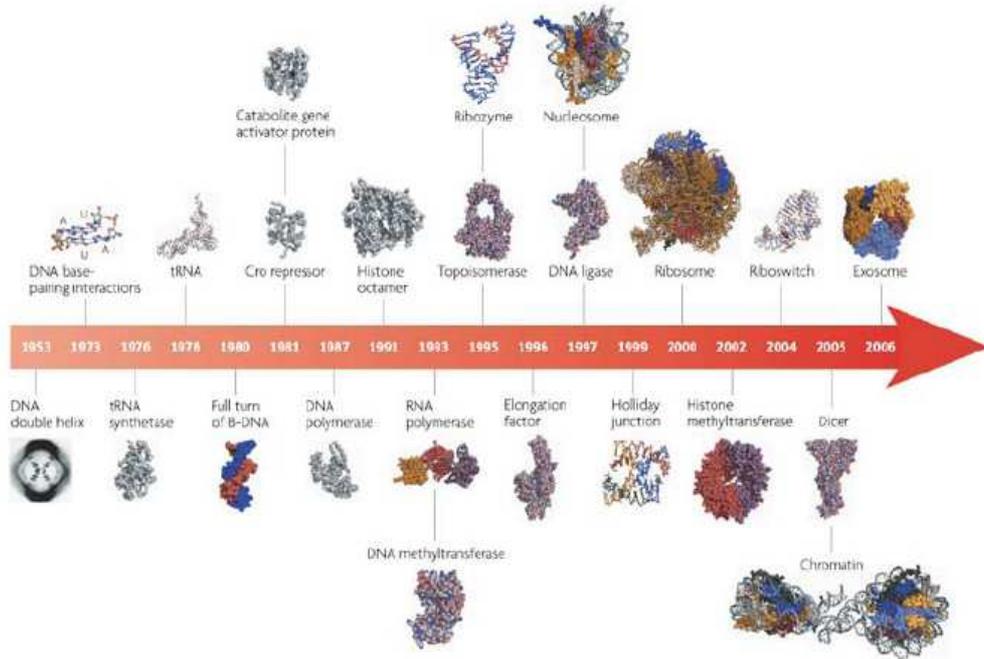
ويستمد البنك معطياته من مخابر مختلف بقاع العالم كالولايات المتحدة، أوروبا وآسيا. كما تم إنشاء موقع إلكتروني لهذا البنك على شبكة الإنترنت (www.rcsb.org/pdb/home) عام 2003 ليوفر أرشيفا إلكترونيًا مجانيًا على شبكة الإنترنت لبيانات الجزيئات الحيوية ثلاثية الأبعاد المحددة تجريبيًا (Berman et al., 2003).



الشكل (2): نافذة العرض لموقع بنك معطيات البروتينات PDB على شبكة الإنترنت.



الشكل (3): نسبة تزايد عدد البروتينات في بنك PDB ، العدد المحقق سنويا (أزرق) العدد الإجمالي (احمر).



الشكل (4): التطور في تعقيد نوع البنيات الثلاثية الإبعاد المحددة تجريبيا للجزيئات الحيوية منذ 1953 إلى غاية 2006.

2.4.1_ بنك المعطيات NCBI

أنشأ المركز الوطني للمعلومات الخاصة بالتقانة الحيوية NCBI التابع للمعاهد الوطنية للصحة عام 1988م وذلك بغرض تطوير برامج وأنظمة خاصة بالجزئيات الحيوية، بالإضافة إلى الحفاظ على معطيات التتابعات الجينية للأحماض النووية التي تم تحديدها من طرف المختصين في هذا المجال، حيث يعتمد بنك NCBI على أنظمة حاسوبية جد متطورة لتخزين و استرجاع وتحليل هذه المعطيات البيولوجية (David L et al., 2007)، كما يتوفر هذا البنك على موقع على شبكة الإنترنت www.ncbi.nlm.nih.gov يمكن من خلاله التحميل المجاني للتتابعات الجينية أو التتابعات الأمينية للبروتينات التي نهدف إلى دراستها.

5.1_الطرائق التجريبية المستعملة في تحديد بنية البروتين

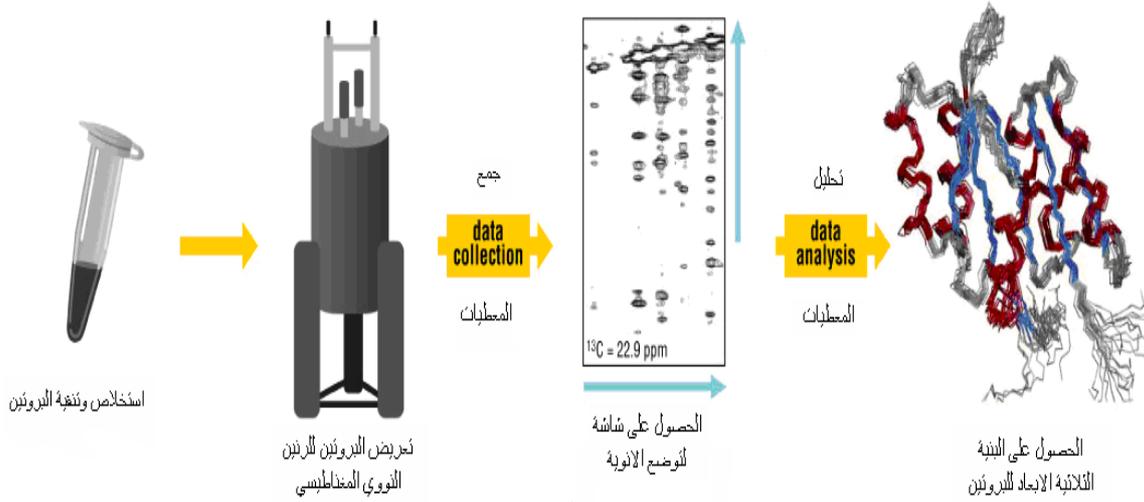
إن فهم العلاقة التي تربط بين بنية ووظيفة الجزئيات الحيوية هو جوهر ما تصبو إليه الفيزياء الحيوية الجزيئية، ومن بين الطرق التجريبية الأكثر استعمالا وانتشارا في هذا المجال التي تصف الجزء الأول من المعادلة بنية-وظيفة هي الدراسة البلورية بواسطة الأشعة السينية X-ray crystallography، والدراسة البلورية بواسطة الرنين النووي المغناطيسي NMR لتحديد البنية ثلاثية الأبعاد للجزئيات الحيوية.

إن التطور الكبير الذي شهدته برامج ومشاريع تحليل تتابعات الحمض الريبي النووي منقوص الأكسجين DNA (Bernal et al., 2001) أدى بدوره إلى تضخم كبير جدا في مخزون بنوك المعطيات الخاصة بالتتابعات النيكلوتيدية والأمينية والبنوية المحددة. ولأن معرفة البنية ثلاثية الأبعاد للبروتين شيء ضروري وأساسي لتحليل الخصائص الوظيفية للبروتين على المستوى الجزيئي وفهم تفاعلاته مع مكونات الخلية الأخرى، حيث تم تحديد أول بنية ثلاثية الأبعاد للبروتينات تجريبيا وهي بنية الميوجلوبين Myoglobin (Kendrew et al., 1958) ثم تلتها بنية الهيموجلوبين Hemoglobin (Perutz et al., 1960) بواسطة الأشعة X.

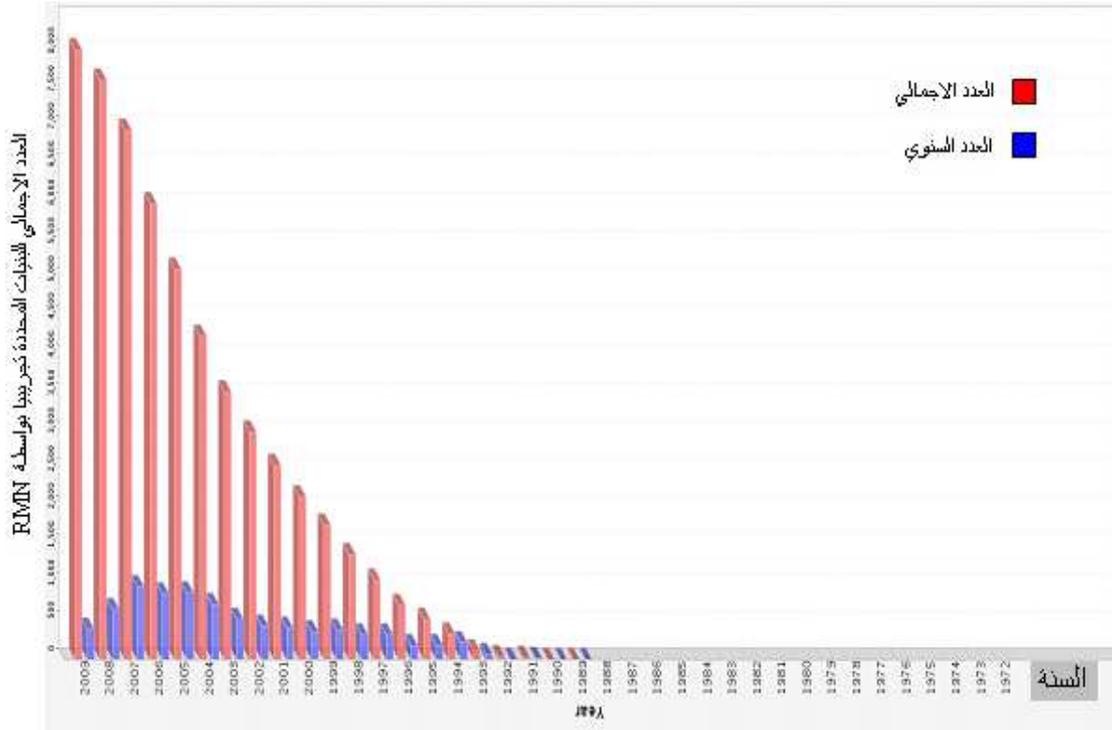
تبقى عملية تحديد البنية ثلاثية الأبعاد تجريبيا صعبة و مكلفة و تستغرق وقتا طويلا مما حد من انتشارها و تطبيقهما عمليا، وزاد في الهوة بين العدد الكبير جدا للتتابعات النيكلوتيدية والأمينية المحددة تجريبيا و النسبة الضعيفة جدا من التتابعات التي تم تحديد بنيتها الثلاثية الأبعاد تجريبيا (Lattman et al., 2004) ما جعل الباحثين يسعون للبحث عن بدائل لطريقتي الرنين النووي المغناطيسي NMR والأشعة X كالاتتماد على الطرق البيومعلوماتية للتنبؤ بالبنية الثلاثية انطلاقا من التتابع الأميني بهدف تقليص الفجوة بين عدد التتابعات المحددة تجريبيا وعدد البنات ثلاثية الأبعاد.

1.5.1-الرنين المغناطيسي النووي «NMR» Nuclear Magnetic Resonance

ظهرت هذه التقنية في البداية كطريقة فيزيائية لتشخيص الخصائص المغناطيسية لبعض الذرات، ثم وجد فيها الكيميائيون أهمية كبرى لأنها توضح العلاقة بين تردد الرنين المغناطيسي للأنوية وطبيعتها، ثم أستغلت فيما بعد من طرف البيولوجيون في تحديد بنية الجزيئات الحيوية حيث أن التطور الكبير من الدراسات الخاصة في هذا المجال، تضمن بصورة كبيرة دراسة البروتينات متبوعة بالأحماض النووية ثم الكربوهيدرات، والأقسام الأخرى من الجزيئات الحيوية (1993 Laura et al.,). ولتقنية الرنين النووي المغناطيسي أهمية كبيرة في عالم البيولوجيا لأنها تسمح بالحصول على البنية ثلاثية الأبعاد للجزيئات الحيوية، حيث يمكن أن تُنفذ هذه التقنية في وسط سائل (وسط محلول مشابه للوسط الفيزيولوجي الطبيعي للبروتينات)، لكن من سلبيات هذه الطريقة أنها خاصة بالجزيئات الصغيرة حيث لا يمكن تطبيقها على البروتينات ذات الوزن الجزيئي المتوسط والكبير.



الشكل(5): يبين مراحل عملية تحديد بنية البروتين بواسطة الرنين النووي المغناطيسي NMR.

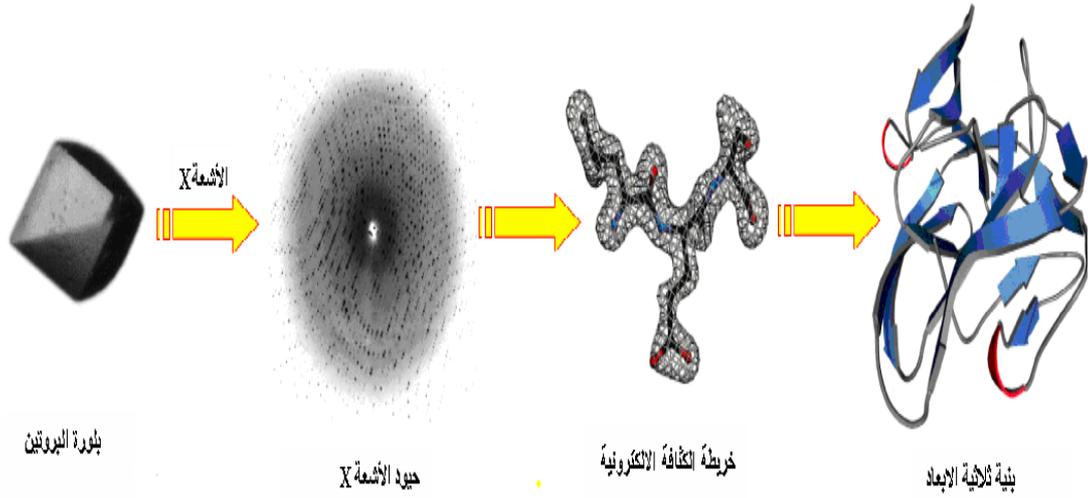


الشكل (6): نسبة تزايد عدد البروتينات في بنك PDB المحددة بواسطة RMN، العدد المحقق سنويا (أزرق) العدد الإجمالي (احمر).

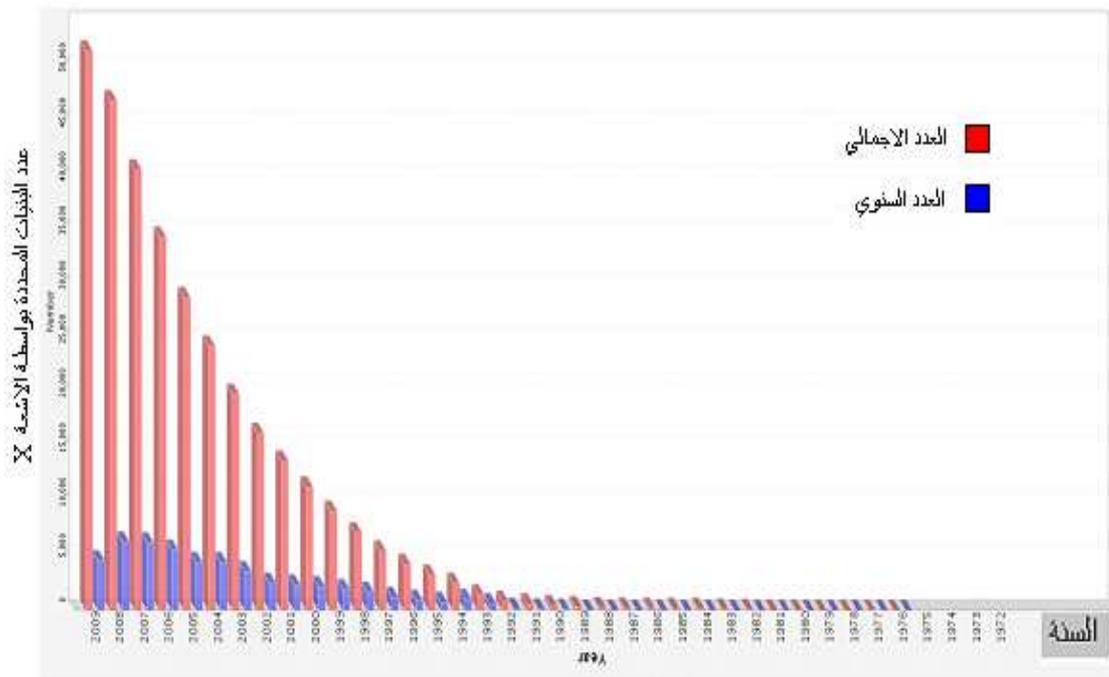
2.5.1_الدراسة البلورية بواسطة حيود الأشعة السينية X ray

تستخدم الأشعة السينية في عدة مجالات وأهمها المجال الطبي والبيولوجي، وذلك بتحليل ترتيب الذرات في أنواع كثيرة من المواد، وخاصة الجزيئات الحيوية، حيث تنتظم الذرات في البلورات على شكل مستويات تفصل بينها مسافات منتظمة، وعندما يسقط شعاع من الأشعة السينية على البلورة، فإن مستويات الذرات تعمل كمرايا صغيرة تحيد الأشعة على نمط نظامي، وكل نوع من البلورات له نمط حيود مختلف عن الآخر. يستخدم العلماء الأشعة السينية لتحليل تركيب وتكوين مواد كيميائية معقدة كثيرة مثل الإنزيمات والبروتينات والأحماض النووية الريبية (DNA، RNA)...الخ.

حيث تصدر المواد أشعة ذات أطوال موجية خاصة بها عندما تتعرض لإشعاع من إلكترونات أو بروتونات عالية الطاقة أو للأشعة السينية X ray (Smyth et al., 2000)، وتسمى هذه الطريقة لتحليل المواد قياس الطيف بالأشعة السينية، ولقد أدت هذه التقنية إلى توصل العلماء إلى اكتشاف بعض العناصر الكيميائية الجديدة.



الشكل (7): يبين مراحل عملية تحديد بنية البروتين بواسطة الأشعة X.



الشكل (8): نسبة تزايد عدد البروتينات في بنك PDB المحددة بواسطة الأشعة X. العدد المحقق سنويا (أزرق) و العدد الإجمالي (احمر).

6.1_ ملفات بنية البروتينات

بنيات الجزيئات الحيوية ثلاثية الأبعاد الموجودة في بنك المعطيات PDB هي عبارة عن ملفات تعريف من الشكل الإلكتروني pdb، حيث يتكون هذا الملف من مجموعة من المعلومات الخاصة بالبنية المدروسة التي تصف تناسق الذرات و الخصائص الكيميائية والبيوكيميائية لها، التفاصيل التجريبية للطريقة التي تم بها تحديد البنية، وكذا وصف بعض الخصائص الهيكلية للبنية مثل الروابط الهيدروجينية ونوع البنيات الثانوية (Qing z et al., 2004). توجد أشكال أخرى من ملفات التعريف الإلكترونية التي يمكن أن تحفظ بها البنيات البروتينية كالمol2 و MmCIF (Westbrook et al., 2006).

```
HEADER      OXIDOREDUCTASE                      03-AUG-05   2AKJ
TITLE       STRUCTURE OF SPINACH NITRITE REDUCTASE
COMPND      MOL_ID: 1;
COMPND      2 MOLECULE: FERREDOXIN--NITRITE REDUCTASE, CHLOROPLAST;
COMPND      3 CHAIN: A;
COMPND      4 EC: 1.7.7.1;
COMPND      5 ENGINEERED: YES
SOURCE      MOL_ID: 1;
SOURCE      2 ORGANISM_SCIENTIFIC: SPINACIA OLERACEA;
SOURCE      3 ORGANISM_COMMON: SPINACH;
SOURCE      4 GENE: NIR;
SOURCE      5 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE      6 EXPRESSION_SYSTEM_COMMON: BACTERIA;
SOURCE      7 EXPRESSION_SYSTEM_VECTOR_TYPE: PLASMID;
SOURCE      8 EXPRESSION_SYSTEM_PLASMID: PET30B;
SOURCE      9 OTHER_DETAILS: SPINACH NITRITE REDUCTASE WAS PRODUCED AS A
SOURCE      10 HIS-TAGGED RECOMBINANT PROTEIN EXPRESSED FROM THE PET30B-
SOURCE      11 NIR PLASMID IN E. COLI. THE E. COLI CELLS WERE CO-
SOURCE      12 TRANSFORMED WITH THE PLASMID PKK233-CYSG THAT CONTAINS THE
SOURCE      13 CYSG GENE ENCODING THE ENZYME THAT CATALYZES THE RATE-
SOURCE      14 LIMITING STEP IN SIROHEME BIOSYNTHESIS.
KEYWDS      X-RAY CRYSTALLOGRAPHY; HEME; ELECTRON TRANSPORT
EXPDTA      X-RAY DIFFRACTION
AUTHOR      U.SWAMY,M.WANG,J.N.TRIPATHY,S.-K.KIM,M.HIRASAWA,D.B.KNAFF,
AUTHOR      2 J.P.ALLEN
REVDAT      1 24-JAN-06 2AKJ 0
JRNL        AUTH U.SWAMY,M.WANG,J.N.TRIPATHY,S.K.KIM,M.HIRASAWA,
JRNL        AUTH 2 D.B.KNAFF,J.P.ALLEN
JRNL        TITL STRUCTURE OF SPINACH NITRITE REDUCTASE.
```

الشكل (9): عرض لبعض محتويات الملف الإلكتروني من نوع PDB للبروتين بواسطة WordPad.

7.1_ البرامج المستعملة في عرض ملفات البروتينات

توجد عدة برامج مستعملة لعرض البنيات ثلاثية الأبعاد للجزيئات الحيوية، منها البسيطة الموجهة لغرض التدريس مثل برنامج Rastop، ومنها ما هو مستعمل لغرض البحث مثل VMD، GenMol، Pymol (Humphrey et al., 1996). ولهذه البرامج أهمية كبيرة في إيضاح كل من التوضع الفراغي والتركييب الكيميائي لبنية البروتين، مواد الارتباط و كيفية الارتباط مما يسهل من فهم آلية عملها.

2_ البروتينات

تعتبر البروتينات من أهم المكونات التي تدخل في تركيب الكائن الحي، فهي جزيئات عضوية حيوية ذات وزن جزيئي كبير تحتل النسبة والأهمية الأكبر بعد الماء في تكوين الكائن الحي حيث تدخل بنسبة أكثر من 50% من الكتلة الجافة للكائن الحي.

اكتشفت البروتينات عام 1839 من طرف العالم Mulder (Mulder et al., 1839) والتي أطلق عليها الاسم اليوناني Proteios والتي تعني الشيء الأول من حيث الأهمية، حيث تضمن البروتينات بمختلف أنواعها أغلب الوظائف الحيوية الخلوية كإنزيمات، نواقل وهرمونات و مستقبلات. يمكن أن تكون عبارة عن سموم أو أجسام مضادة تدخل في استتساخ المادة الوراثية كما تشترك في التركيب الهيكلي أو البنية الخلوية بصفة عامة، كل هذه الوظائف تكون بصورة عامة على شكل تثبيث البروتين بجزء آخر (والذي يمكن أن يكون أيضا بروتين) هذا التثبيث يكون نوعيا وتلعب السلاسل الجانبية للأحماض الأمينية الدور الأكبر فيه.

تتركب البروتينات من تسلسل مضبوط ترتيبا و نوعا وعددا من الأحماض الأمينية المشفرة من طرف الجينات يدعى هذا التسلسل بالتتابع الأميني للبروتين أو البنية الأولية للبروتين التي يمكن أن تتطوي وتأخذ أشكال منتظمة ومضبوطة تسمى البنية الثانوية (حلزون ألفا، أوراق بيتا)، التوضع الفراغي للبنيات الثانوية بعد عملية الانطواء ينتج عنه البنية الثلاثية التي هي الدعامة الوظيفية للبروتين.

1.2_ العلاقة بين بنية ووظيفة البروتينات

وظيفة البروتين ترتبط كليا ببنيته الثلاثية الأبعاد، حيث لا يمكن فهم آلية عمل البروتين إلا بعد تحديد ومعرفة بنيته، ففي عام 1954 بين العالم Anfinsen (جائزة نوبل البيوكيمياء عام 1972) أن المعلومات الضرورية لانطواء البروتين إلى الشكل النهائي والوظيفي موجودة في بنيته الأولية أي في تتابع أحماضه الأمينية (Anfinsen et al., 1973)، أطلق على هذه الملاحظة مبدأ Anfinsen. نظريا يمكن لبروتين واحد أن يأخذ عدة أشكال من الانطواء لكن أغلب البروتينات تأخذ تلقائيا انطواء واحدا مستقرا وخصوصا، هذا الانطواء الخاص يرجع إلى تفاعل مجموعات الهيكل الكربوني وكذا مجموعات السلاسل الجانبية للأحماض الأمينية فيما بينها ومع الماء حيث نجد بعض التوضعات مستقرة والبعض الآخر أقل استقرارا (Alberts et al., 1994)، حيث أحسن مثال لتبيان العلاقة بين تتابع الأحماض الأمينية للبروتين وبنيته ثلاثية الأبعاد مأخوذ من دراسة Christian Anfinsen لإنزيم الريبونوكلياز Ribonuclease (Anfinsen et al., 1973).

بين Anfinsen تجريبيا أن البروتين الذي له انطواء عشوائي يصبح غير نشط وبذلك يفقد وظيفته. يمكن إيجاد بنية البروتين الأصلية (الأولية) عندما تكون معزولة في محلول وذلك بعد إزالة الانطواء الطبيعي لها بإضافة مواد مزيلة للانطواء. ويمكن القول أن المعلومات اللازمة عن انطواء أي بروتين تُستمد أساسا من تتابع أحماضه الأمينية (Alberts et al., 1994).

وعلى اعتبار أن كل حمض أميني له فقط ثلاث حالات من التوضع للزاويتين (ψ و ϕ) الممكنة، فإنه إن افترضنا أن بروتين ما يحتوي على 100 حمض أميني، فإنه نظريا بإمكانه أن ينتج عنه 100^3 شكل انطوائي ممكن، أما في الوسط الخلوي فإن البروتين ينطوي إنطواء واحدًا وفي زمن قصير جدا يقدر بأجزاء الثانية إلى الدقيقة (Dinner et al., 1996).

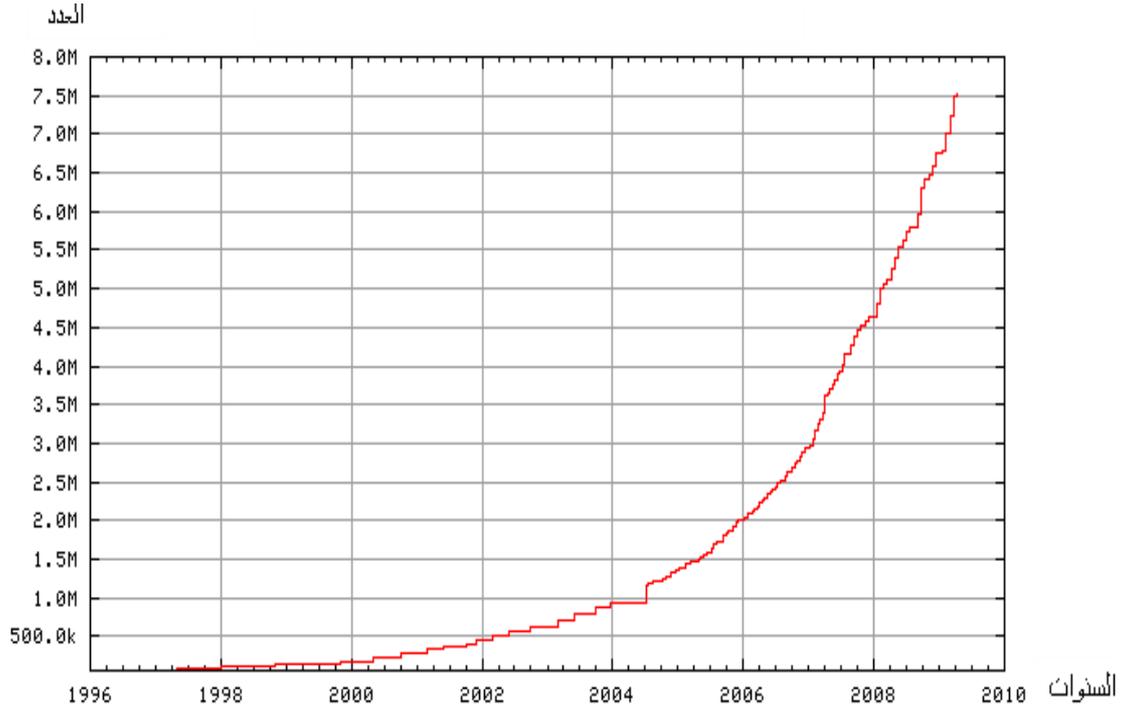
ولهذا فإن عدد كبير من الأمراض لها علاقة مباشرة بالانطواء الخاطئ للبنية ثلاثية الأبعاد للبروتين الذي يؤدي إلى فقدان جزئي أو كلي للوظيفة، وفي بعض الأحيان اكتساب وظيفة جديدة، كأمراض النكسة العصبية Neuro-degenerative مثل مرض الزهايمر Alzheimer ومرض باركينسون Parkinson (Dobson et al., 2004)، حيث أن التغير في تتابع أو نوع أو عدد الأحماض الأمينية هو السبب الرئيس في تغير البنية ثلاثية الأبعاد للبروتين، هذا ما جعل التحليل البنيوي الدقيق للبروتينات يحوز أهمية بالغة في عالم الأمراض خاصة والبيولوجيا عموما.

انطلاقا من مبدأ Anfinsen (Anfinsen et al., 1952) فإنه من الممكن نظريا الوصول إلى البنية ثلاثية الأبعاد لأي بروتين انطلاقا من تتابعه الأميني فقط، هذا ما قاد إلى ظهور عملية التنبؤ ببنية البروتين Protein Structure Prediction، حيث نميز إستراتيجيتين أو طريقتين للتنبؤ بالبنية ثلاثية الأبعاد لأي بروتين وهي النمذجة بالتماثل Homology modeling و التنبؤ من جديد Ab initio (Gibas et al., 2002).

3_التنبؤ ببنية البروتين Protein Structure Prediction

تحديد البنية ثلاثية الأبعاد للبروتينات هي المفتاح الأساسي لفهم وظيفتها والتي تعتبر من المسائل والتحديات الأساسية التي تواجهها البيولوجيا الجزيئية، والطريق الأمثل الذي استخدم في البداية لاستخراج ومعرفة المعلومات الخاصة بالبروتينات هي تحديد بنيتها ثلاثية الأبعاد تجريبيا بواسطة الرنين النووي المغناطيسي أو الأشعة X لكن هاتين الطريقتين ورغم نجاعتهما وجودة نتائجها إلا أنهما لم يحلا مشكل تحديد البنية البروتينية كليا لعدة أسباب حسب (et al., 2000): (Jones):

- 1- طريقتي الرنين النووي المغناطيسي والأشعة X متوفرة في عدد قليل جدا من المخابر لتكلفتها الباهظة.
- 2- بعض البروتينات يصعب تحديد بنيتها البلورية بسهولة نظرا لصعوبة الاستخلاص والتقية.
- 3- تحديد البنية البلورية لبروتين واحد تجريبيا بواسطة الأشعة X يستغرق وقتا طويلا من عدة شهور إلى عدة سنوات.
- 4- تحديد البنية ثلاثية الأبعاد للبروتينات بواسطة الرنين النووي المغناطيسي وإن كانت أسرع من الأشعة X إلا أنه لا يمكن تطبيقها حاليا إلا على البروتينات التي لها عدد أحماض أمينية اقل من 100 حمض أميني.
- 5- هناك فجوة كبيرة جدا بين العدد الكبير للتتابعات البروتينية المحددة والمقدرة (14-Apr-2009) بحوالي 7527796 تتابع أميني (www.ebi.ac.uk/uniprot)، والعدد القليل نسبيا من البنيات ثلاثية الأبعاد للبروتينات التي تم تحديدها تجريبيا والتي تقدر بحوالي 57133 بنية (http://www.rcsb.org/pdb/static).
- 6- كنتيجة لمشاريع تحديد تتابع الجينوم فإن التتابعات الأمينية للبروتينات التي يتم تحديدها تتزايد بشكل كبير جدا.
- 7- إن افترضنا أنه سيتم تحديد كل البنيات ثلاثية الأبعاد لكل البروتينات المحددة التابع تجريبيا بواسطة الرنين النووي المغناطيسي والأشعة X فإنها تتطلب أكثر من 500 عام .



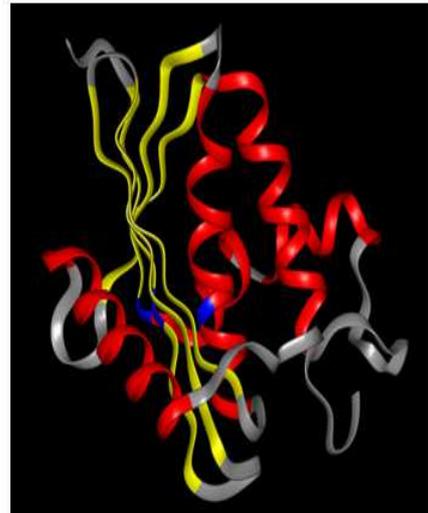
الشكل (10): يمثل تطور عدد التتابعات الأمينية للبروتينات في بنك Uniprot.

ونتيجة للأسباب والعوائق السابقة، طوّر العلماء برامج حاسوبية تعتمد على الخوارزميات الرياضية القادرة على رسم والتنبؤ بالبنية ثلاثية الأبعاد للبروتين والتي تعرف ببرامج التنبؤ بالبنية ثلاثية الأبعاد Programs Protein Structure Prediction.

```

...GCGGGACTCAAGAGTAGCCTTCCTCGAGGACCTGC
CTTTCCCATTTGCTGCCTGAAAGTTAATGTTTCTTGCTG
GCCAAATCAGGGACATGCCCGCATTAGCGGGATGAGTG
GGTGTTCGGGCAGGGATGTGGTCATTGACGGCCAGTGA
GGGCGAGAGTACCACGGCCCACTTCTCCTTGGAGCTG
GAGATGAGGGGCTGGGCACCCGTGGAATAGGCATGAGG
CCAGAAGAGAGTGACAGCGAGCTCCTTGAGGATGAGGA
GGATGAAGTGCCTCCTGAACTCAGATCATTGTTGGCA
TCTGTGCCATGACCAAGAAATCCAAGTCCAAGCCAATG
ACTCAAACTCCTAGAGCGACTCTGCAGATTTGACTACCT
GACTGTTGTCAATTCTGGGAGAAGATGTAATCCTTAATG
AACCTGTGGAAAACCTGGCCATCCTGCCACTGCCTCATC
TCTTTCCACTCCAAGGCTTTCCCTCTGGACAAAAGCTGT
TGCTTACTCCAAGCTTCGAAAACCCCTTTCTTATCAATG
ATCTGGCCATGCAGTATTACATCCAAGATAGGAGGGAG
GTGTACCAGATCCTGCAGGAAGAGGGTATTGATCTGCC
TCGATATGGCTGTGCTCAACCGTGATCCTGCCCGGCTG
AGGAATGCAACCTGATAGAGGTGAAGACCAAGTAGAG
GTCAATGGAGCTGTCTTTCCGATCCTGCAGGAAGCT...

```



الشكل (11): يوضح مخطط التنبؤ بالبنية ثلاثية الأبعاد انطلاقاً من التتابع الأميني.

إن التنبؤ ببنية البروتين من أقدم المسائل التي واجهت البيومعلوماتية المختصة بدراسة البنية. و قد استعملت عدة طرق مختلفة للتغلب على هذا المشكل على مدى السنين، وكانت أكثر هذه الطرق نجاحا تلك التي اعتمدت على المعلومات الخاصة بالعلاقة بين التتابع الأميني و البنية الثلاثية (Sali et al., 1995; Moult et al., 2005)، حيث تعمل هذه الطريقة على بناء نموذج نظري للبنية ثلاثية الأبعاد للبروتين انطلاقاً من المعطيات الأولية لتتابعه الأميني والتي تدعى بطريقة النمذجة بالتماثل Homology modeling، أين تنطلق مباشرة من التتابع الأميني للبروتين للتنبؤ ببنيته الثلاثية وهذه الطريقة هي التي سنعتمد عليها لبناء النموذج النظري للبنية ثلاثية الأبعاد لإنزيم Ferredoxin-nitrite reductase عند نبات *Arabidopsis thaliana*.

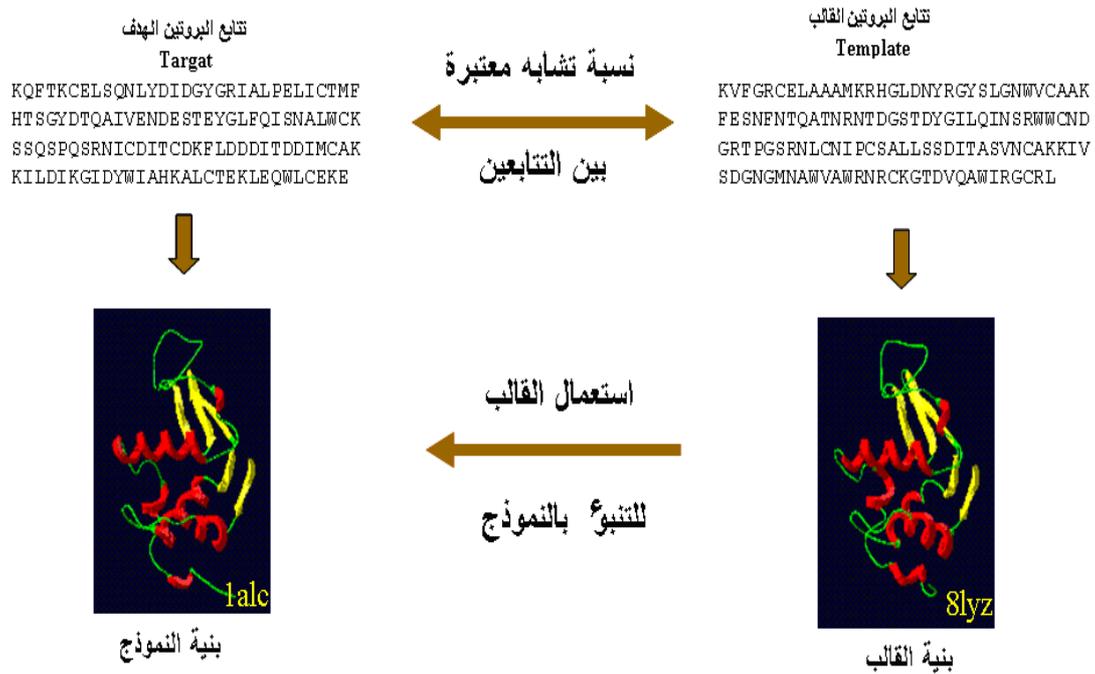
1.3_النمذجة بالتماثل Homology modeling

إن عدد البروتينات المحددة البنية تجريبيا الآن تفوق 50000 بروتين في بنك معطيات البروتينات PDB وهي في تزايد مستمر بمقدار حوالي 6000 بنية جديدة سنويا، والوسائل التجريبية سواء كانت حيود الأشعة السينية X أو الرنين النووي المغناطيسي NMR تبقى بطيئة التنفيذ ومحصورة على نطاق ضيق، مقارنة بطرق ووسائل تحديد التتابع الأمينية (مشاكل في الاستخلاص، الذوبان، تنقية وبلورة البروتين)، كما أن تحديد بنية كل البروتينات المستخلصة تجريبيا شيء مستحيل، وهذا نظرا إلى العدد الكبير جدا للجينومات والجينات المحددة (Pearl et al., 2005)، لذا لجأ الباحثون إلى طريقة النمذجة بالتماثل Homology modeling.

تعتبر طريقة النمذجة بالتماثل Homology modeling من أقدم طرق التنبؤ ببنية البروتينات، وقد تطورت بشكل سريع خلال العشرين سنة الماضية (Tramontano et al., 2003). وتعتمد هذه الطريقة أساسا على المبدأ الذي ينص على أن التتابعات الأمينية المتماثلة من وجهة نظر تطويرية يكون لها بنيات ثلاثية الأبعاد متشابهة الانطواء (Holm et al., 1996)، أي أنه انطلاقا من التشابه في التتابع الأميني يفترض التشابه من حيث البنية الثلاثية

الأبعاد. وعليه فإن النمذجة بالتماثل تسمح بالحصول على نموذج نظري ثلاثي الأبعاد لبروتين لم تحدد بنيته تجريبيا (Sali et al., 1993).

إن نجاح ومصداقية طريقة النمذجة بالتماثل متعلق بشكل أساسي ومباشر بنسبة التطابق بين تتابع البروتين الهدف والبروتين المحدد البنية تجريبيا و الذي سيستعمل كقالب Template لبناء النموذج النظري الذي نسميه بالبروتين الهدف Target. بحيث إذا كانت نسبة التطابق بينهما أكثر من 50 % يمكن اعتبار أن التتابعين متطابقين ويكون التنبؤ عالي النوعية. كما أن النماذج المتنبأ بها توازي البنيات المحددة تجريبيا ذات الدقة العالية أي ذات القيمة المنخفضة جدا لدقة التصميم Resolution (Kopp et al., 2004). أما إذا كانت نسبة التطابق تتراوح بين 30% و50% فإن دقة البنيات المتنبأ بها توازي بنيات محددة تجريبيا لها دقة تصميم أقل من 3,5Å (Kopp et al., 2004). وإذا كانت نسبة التطابق أقل من 30% فإن هناك احتمال كبير بأن النموذج النظري البروتيني الذي سنحصل عليه يحمل أخطاء كثيرة (Vitkup et al., 2001;Kopp et al., 2004).



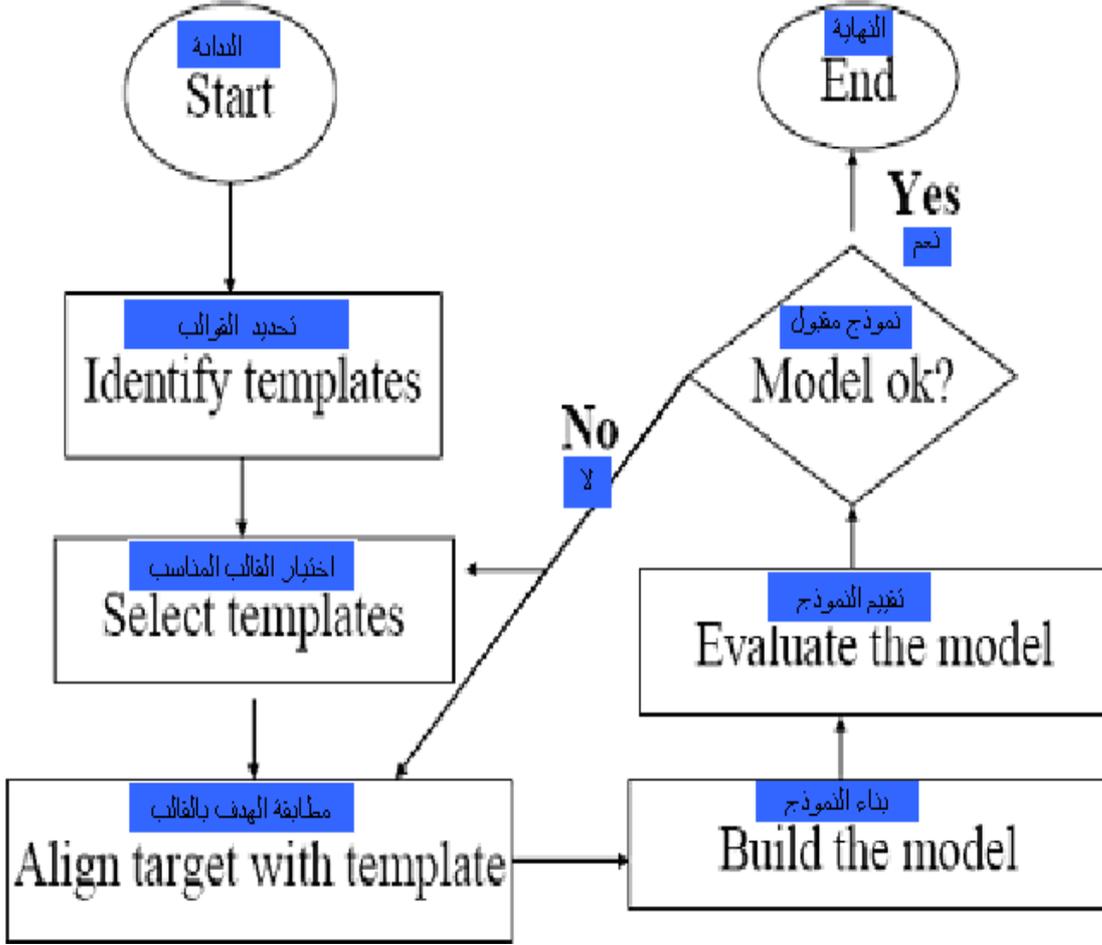
الشكل (12): يوضح مخطط عام للتنبؤ بالبنية ثلاثية الأبعاد للبروتينات عن طريق النمذجة بالتماثل

.Homology modeling

إن الهدف من النمذجة بالتماثل هو بناء نموذج نظري ثلاثي الأبعاد لبروتين غير محدد البنية تجريبيا والذي نسميه بالبروتين الهدف Target، وذلك بالاعتماد على تتابعات بروتينات مشابهة

له من حيث التتابع ومحددة البنية ثلاثية الأبعاد (تجريبيا) والتي نسميها بالبروتين قالب له من حيث التتابع ومحددة البنية ثلاثية الأبعاد (تجريبيا) والتي نسميها بالبروتين قالب (Bajorath et al., 1994; Blundell et al., 1987; Johnson et al., 1994). وعلى الرغم من التطور الكبير الذي عرفته عملية التنبؤ بالبنية ثلاثية الأبعاد للبروتينات منذ البداية (Koehl et al., 1999)، وتتنوع طرقها تبقى النمذجة بالتماثل الأدق والأكثر استعمالا وانتشارا من بين كل الطرق، وهذا راجع إلى أن دقة ومصداقية هذه الطريقة ممتدة على مجال واسع، ابتداء من النماذج ذات دقة التصميم Resolution أو المعيار الصغير، المتوسط، والكبير والذي سيقارن لغرض تقييمه بمتوسط معيار التصميم المحدد تجريبيا سواء بالرنين النووي المغناطيسي أو بالأشعة X (Sánchez et al., 1997)، حيث كلما كان هذا المعيار صغيرا كلما كانت مصداقية النموذج دقيقة ومفيدة من الناحية البيولوجية.

إن البنية ثلاثية الأبعاد للبروتينات التي تنتمي إلى نفس العائلة تكون غالبا متقاربة ومحفوظة (Lesk et al., 1980)، وعلى هذا الأساس إن كان هناك تشابها بين بروتينين على مستوى التتابع الأميني فإنه من الممكن التنبؤ بالتشابه بينهما على المستوى البنيوي، ورغم ذلك فإنه حتى وإن لم يكن لبروتين ما تشابه واضح مع تتابع البروتينات الأخرى فإنه يمكن أن يكون له ما يشبهه على المستوى البنيوي. وقد قدر أن ما يقارب ثلث التتابع الكلي لأي بروتين له علاقة تشابه على الأقل ببروتين واحد معروف البنية (Fischer et al., 1997; Huynen et al., 1998; Jones et al., 1999; Rychlewski et al., 1998). وهذا نظرا للعدد الكبير للبروتينات معلومة البنية والتي تقدر بأكثر من 500000 بنية بروتينية (Bairoch et al., 1999)، حيث يمكن تطبيق طريقة النمذجة بالتماثل على حوالي 150000 بروتين مجهول البنية اعتمادا على 10000 بروتين فقط محدد البنية تجريبيا (Abola et al., 1987; Berman et al., 2000). علاوة على ذلك فإن فائدة هذه الطريقة في تزايد مستمر وذلك لأن عدد الانطواءات المختلفة للبروتين محدودة (Zhang et al., 1997)، وكذا لزيادة البروتينات التي يتم تحديد بنيتها تجريبيا يوما بعد يوم (Holm et al., 1996).



الشكل (13): مخطط عام لمرحل النمذجة بالتماثل Homology modeling.

2.3_ تطبيقات النمذجة بالتماثل

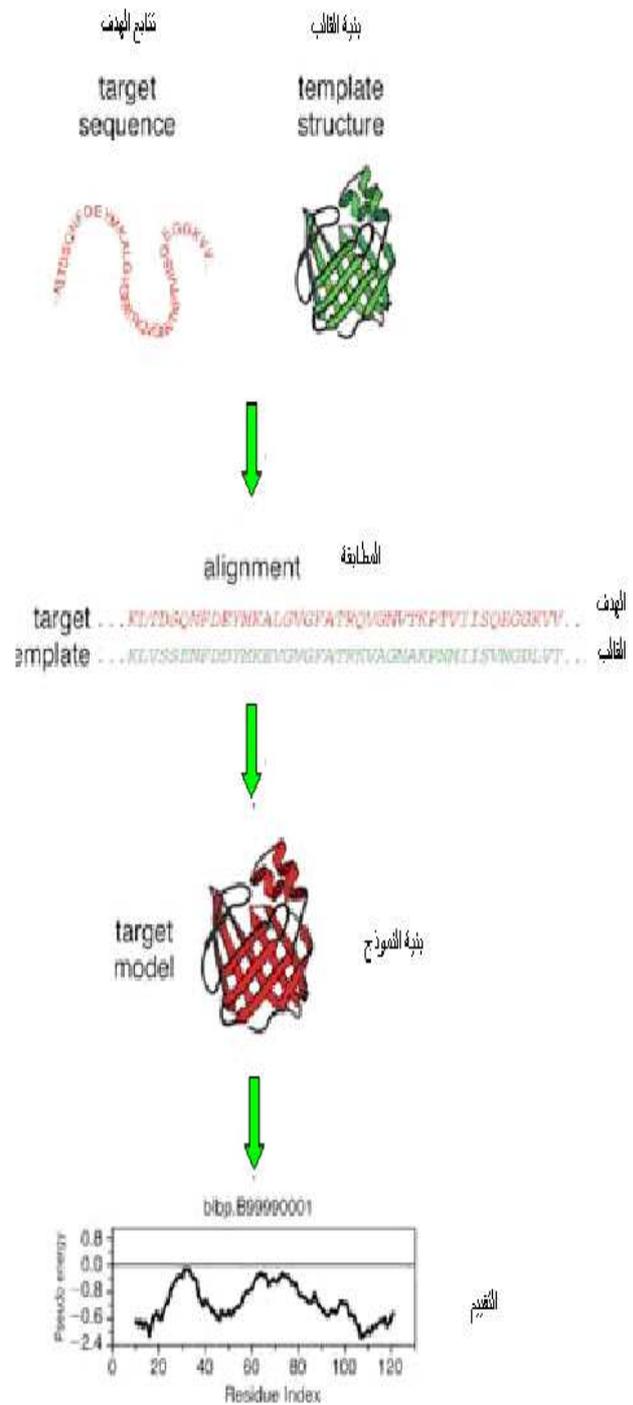
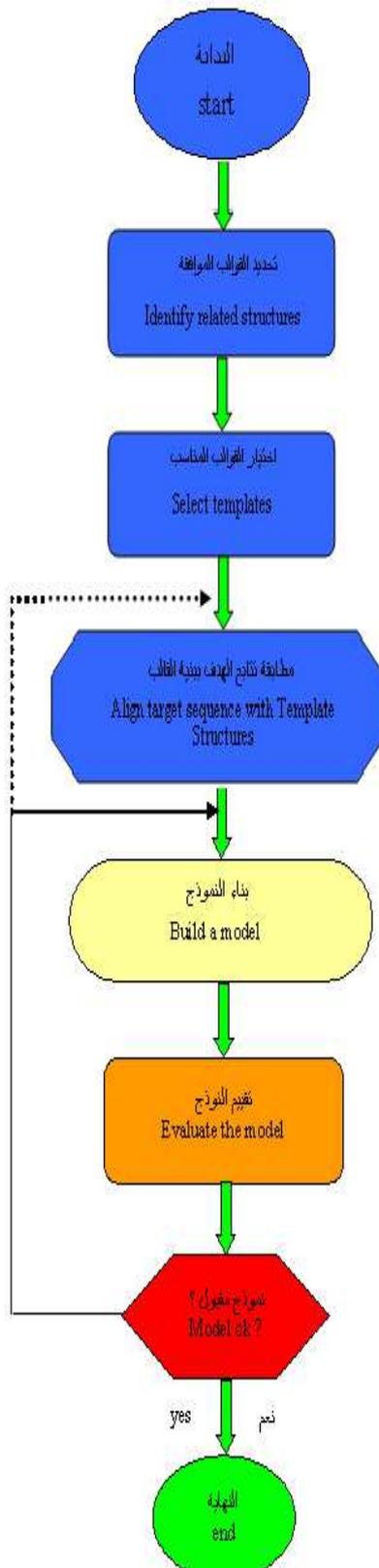
لعبت نمذجة البروتينات دورا هاما ليس فقط في بناء نموذج بنيوي ثلاثي الأبعاد للبروتين وإنما أيضا في تصميم بنية المثبطات Inhibitors وتصميم أيضا مواد الارتباط الخاصة، فعلى سبيل المثال النماذج النظرية لـ HIV-Proteinase ومثبطاته استعملت ودرست لسنوات عديدة قبل تحديد البنية البلورية تجريبيا بواسطة الأشعة X للمعقد إنزيم- مثبط (Weber et al., 1989)، كما هو الحال في نماذج الـ Renin الدقيقة و التي سمحت ببناء مركبات رئيسية ذات شكل حلقي لتقليل مرونتها أو الإنقاص من حجمها لزيادة قابلية امتصاصها المعوي. كما يمكن حذف بعض الروابط الببتيدية لتخفيض التحلل البروتيني أو إضافة مجموعات كيميائية للرفع من الذوبان في الدهون أو تغيير مجموعات كيميائية للرفع من تخصص مادة الارتباط (Plattner et al., 1986).

كما يمكن أن تستغل عملية النمذجة بالتماثل في حقل هندسة البروتينات مثل دراسة المواقع الفعالة ومواقع الارتباط و تقييم مدى ألفة وارتباط مواد الارتباط نظريا مع البروتينات الأصلية والمطفرة قبل تطبيقها تجريبيا، وهذا ما يوفر وسيلة لبناء الجزيئات الكيميائية الممكنة الارتباط والتفاعل وعلى العموم يمكن تلخيص استعمالات النمذجة بالتماثل حسب (Roberto et al., 2000) فيما يلي:

- 1- بناء البنية ثلاثية الأبعاد للبروتينات غير محددة البنية تجريبيا.
- 2 - تصميم وإحداث الطفرات النظرية لاختبار الفرضيات المطروحة حول الوظيفة قبل تطبيقها العملي.
- 3- تحديد كل من موقع الارتباط والموقع الفعال.
- 4- البحث والتنبؤ عن المواد الممكنة للارتباط بالموقع الفعال.
- 5- تصميم وتحسين مواد الارتباط الممكنة.
- 6- دراسة التحام بروتين بروتين.

4_ الخطوات الرئيسية للنمذجة بالتماثل

تتم طريقة النمذجة بالتماثل بشكل عام في أربع خطوات أساسية (S´anchez et al., 1997)، وهي على التوالي: الحصول على التتابع الأميني للبروتين الهدف Target و البحث عن البروتينات القالب Templates الموافقة، ومن ثم اختيار البروتين القالب المحددة بنيته تجريبيا و الأكثر ملائمة مع البروتين الهدف وكذا مطابقة البروتين الهدف ببنية البروتين أو البروتينات القالب الذي تم اختيارها و بناء النموذج النظري للبروتين الهدف. وأخيرا تقييم دقة وصحة النموذج أو النماذج التي تم بناؤها، كما يمكن تكرار بناء النموذج حتى نحصل على نتيجة مرضية أي نموذج مقبول كما هو موضح في المخطط (الشكل14).



الشكل (14): مخطط مفصل لمرحل النمذجة بالتماثل Homology modeling.

5_تمثيل وأهمية النيتروجين في النبات

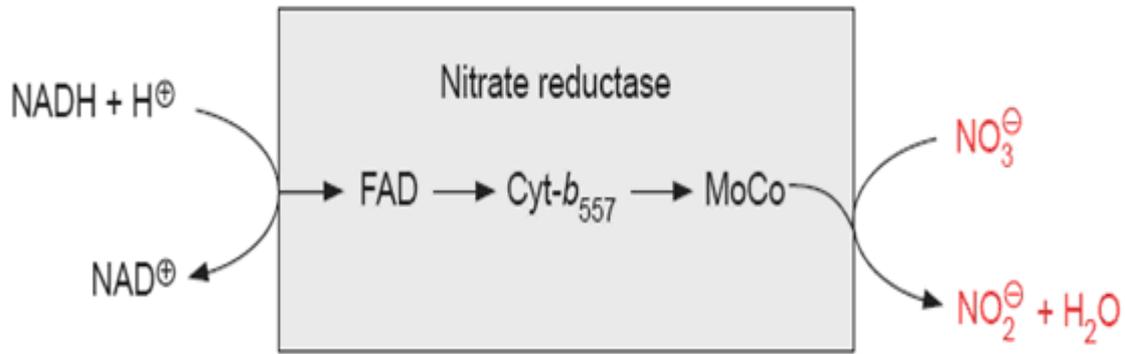
يأتي النيتروجين الرابع في الترتيب بعد الكربون و الهيدروجين والأكسجين في نسبة تركيبه للكائن الحي، إذ يدخل في تركيب مركبات هامة مثل البروتينات والأحماض الأمينية، والأحماض النووية وبعض منظمات النمو وفي الكثير من الفيتامينات، ونظرا لهذه الأهمية الكبيرة فللنيتروجين دور ودخل كبير في جُل وربما كل التفاعلات الكيميائية الحيوية التي تُشكل عصب الحياة، كما يعتبر في غالب الأحيان العنصر المحدد الذي يعيق نمو النبات (Fowden et al., 1979). ولغرض النمو أغلب النباتات تتحصل على النيتروجين بشكل كلي من التربة والذي يتمثل أساسا في الأمونيا NH_4 والنترات NO_3^- حيث تعتبر الأمونيا الشكل النهائي للنيتروجين اللاعضوي في النبات.

إنّ النسبة الكبيرة من النيتروجين وأهميّة هذا العنصر في تركيب النبات وتحولاته الغذائية واحتياج النبات لمدد دائم لهذا العنصر الحيوي جعل منه عنصرا ضروريا، حيث يشكل النيتروجين الجوي حوالي 79% من نسبة غازات الغلاف الجوي ورغم هذه النسبة العالية إلا أنه غير متاح للنبات ولهذا فالكمية الكبرى منه يتم تثبيتها عن طريق تدخل الكائنات الحية الدقيقة التي تعيش في التربة.

تمتص غالبية النباتات الراقية النيتروجين في صورة نترات NO_3^- (Hirsch et al., 1999; Crawford et al., 1998) من التربة ولكن النيتروجين في صورته هذه لا يستغل مباشرة من قبل النبات بل عليه أن يختزل إلى أمونيا قبل أن يدخل في تركيب المركبات النيتروجينية الموجودة في النبات فيما بعد. عملية اختزال النترات NO_3^- إلى الأمونيا NH_4 تتم في مرحلتين:

1.5_المرحلة الأولى:

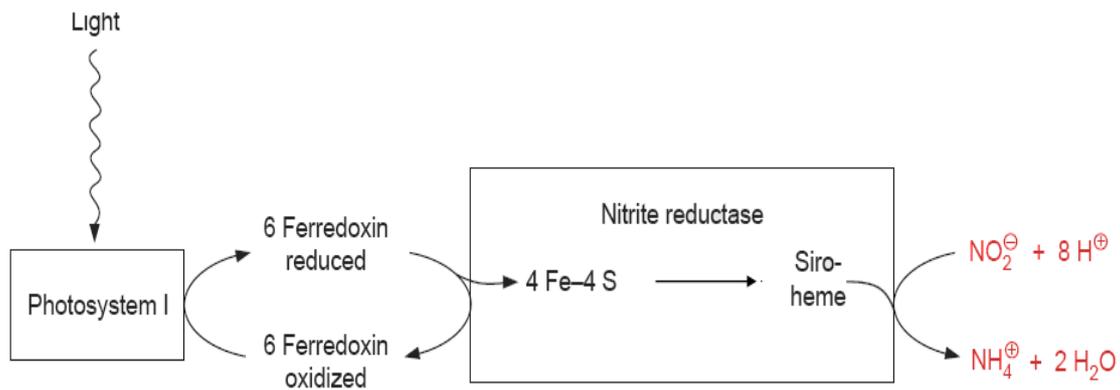
تتم هذه العملية داخل السيتوبلازم، بحيث بعد دخول النترات NO_3^- إلى داخل الخلايا يتم فيها اختزالها إلى نترت (NO_3^- إلى NO_2^-) بواسطة إنزيم سيتوبلازمي وهو الـ Nitrate reductase كما هو موضح في (الشكل 15).



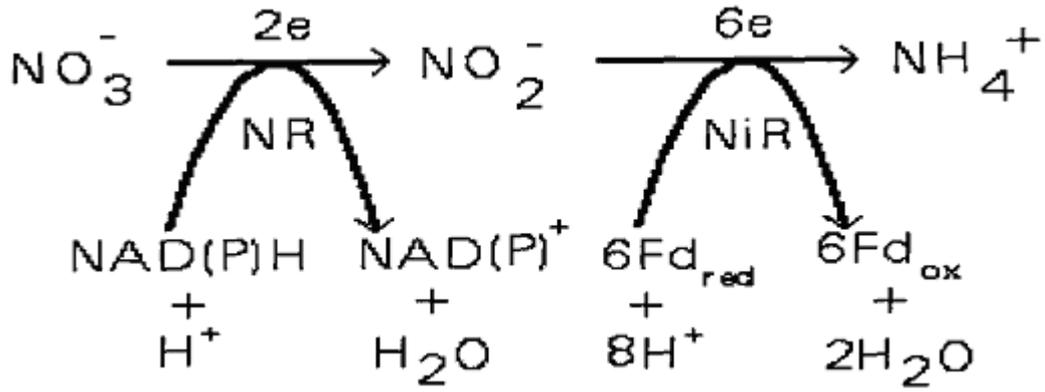
الشكل (15): يمثل المرحلة الأولى وهي إرجاع النترات NO_3^- إلى نترت NO_2^- عند النباتات الراقية بواسطة إنزيم الـ Nitrate reductase (Hans-Walter et al., 2005).

2.5_ المرحلة الثانية

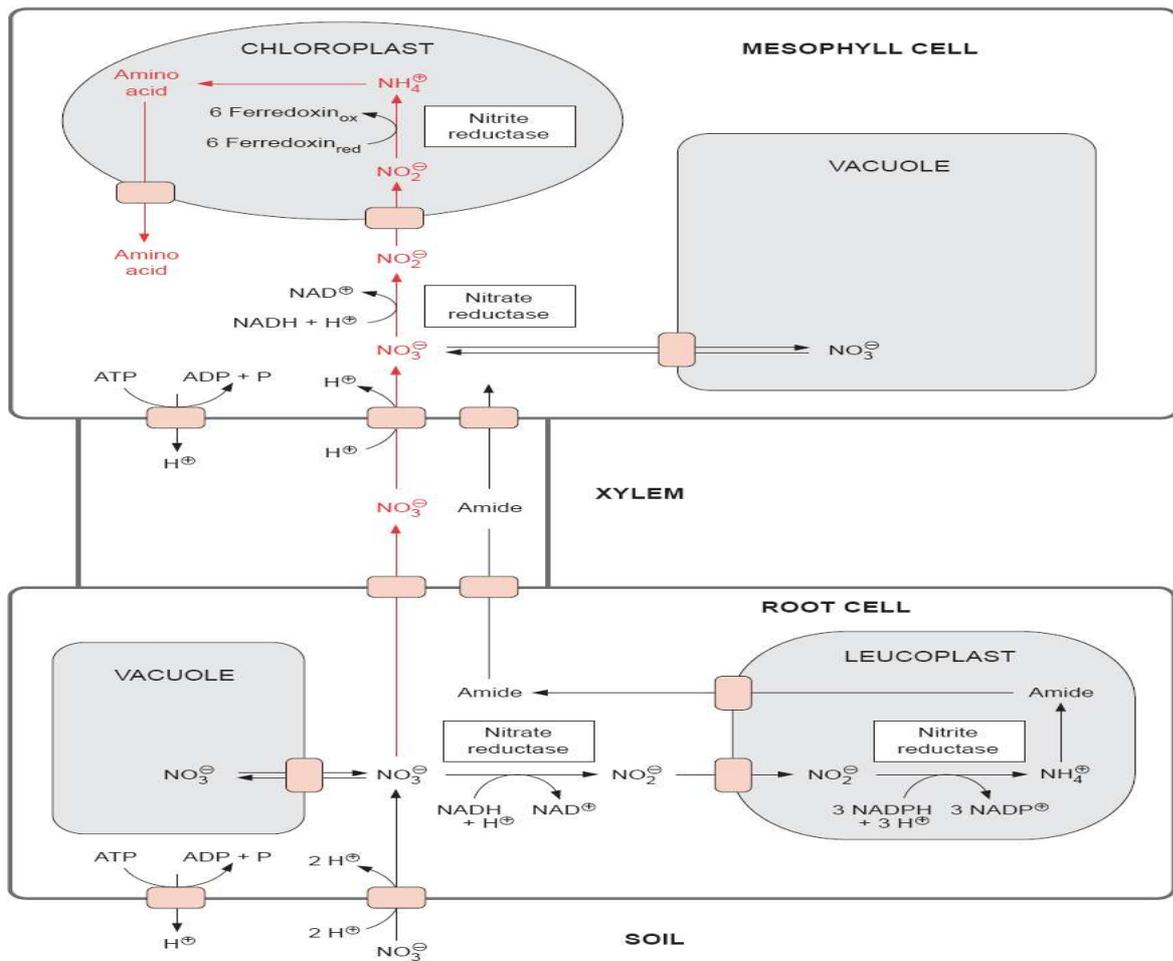
بعد تحويل النترات إلى نترت (التي تعتبر سامة بكميات معينة). تهاجر هذه الأخيرة إلى داخل الصانعات الخضراء في الأنسجة اليخضورية و إلى البروبلاست في الأنسجة غير اليخضورية، أين تحول إلى أمونيا NH_4^+ بواسطة إنزيم Nitrite reductase (René et al., 1998) ، ثم تدمج الأمونيا بعد ذلك في بناء الأحماض الأمينية بواسطة إنزيمي Glutamine synthetase و Glutamate synthase على التوالي كما هو موضح في مخطط (الشكل 16):



الشكل (16): يمثل المرحلة الثانية وهي إرجاع النترت إلى الأمونيا عند النباتات الراقية بواسطة إنزيم Nitrite reductase (Hans-Walter et al., 2005).



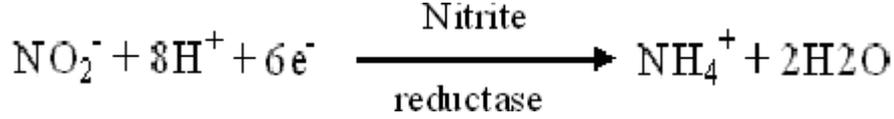
الشكل (17): يمثل المرحلتين الأولى والثانية لإرجاع النترات إلى نترات ثم إلى الأمونيا عند النباتات الراقية.



الشكل (18): مخطط مفصل لتمثيل الأزوت عند النباتات الراقية وعمل وتواجد الإنزيمات المشاركين (Hans-Walter et al., 2005) NIR و NR.

6_ إنزيم الـ Nitrite reductase

يعرف الـ Nitrite reductase النباتي أيضا بالـ Ferredoxin: nitrite reductase والذي يحفز إرجاع النتريت NO_2^- إلى الأمونيا NH_4^+ بواسطة 6 إلكترونات كما هو موضح في المعادلة:



عملية الإرجاع هذه تعتبر الخطوة الأساسية في تمثيل الأزوت عند النباتات الراقية والكائنات المجهرية، التي لها القدرة على القيام بعملية التركيب الضوئي، حيث يعمل الفيريديوكسين Ferredoxin كمانح طبيعي للإلكترونات للـ Nitrite reductase داخل الكلوروبلاست (Knaff et al., 1991).

الـ Ferredoxin:nitrite-reductase هو بروتين يقوم بعمله داخل الصانعة الخضراء لكن يشفر له عن طريق الجينوم النووي أي أنّ مورثته تقع على الحمض النووي الريبي منقوص الأكسجين ADN للنواة (Lopez-Ruiz et al., 1985). كما يتواجد أيضا في خلايا الأنسجة غير اليخضورية مثل الجذور وبالضبط داخل طليعة الصانعة أين يستعمل إنزيم الـ Ferredoxin كمانح طبيعي للإلكترونات (Suzuki et al., 1985). في الأنسجة غير اليخضورية يُرجع الفيريديوكسين Ferredoxin بواسطة NADPH الناتج عن طريق أكسدة بنتوز الفوسفات (Bowsher et al., 1989).

1.6_ مقارنة بين أنواع الـ Nitrite-reductase

الـ Nitrite-reductase (EC 1,7,7,1) عند النبات يتميز عن الـ Nitrite reductase (EC 1,6,6,4) الموجود عند البكتيريا والفطريات. بحيث أنّ الـ Nitrite reductase عند البكتيريا والفطريات يستعمل NADPH كمانح طبيعي للإلكترونات بدلا من الـ Ferredoxin عند النبات، كما أنه عبارة عن بروتين يتكون من وحدتين متجانستين Homodimer بينما الـ Nitrite reductase عند النبات يتكون من سلسلة ببتيدية واحدة Monomeric. كما يحتوي الـ Nitrite reductase للبكتيريا والفطريات على مجموعتين غير بروتينيتين وهما مجموعة الـ Siroheme و مجموعة $[\text{Fe}_4\text{-S}_4]$ واللّتين توجدا أيضا عند الـ Nitrite reductase النباتي، لكن يحتوي الـ Nitrite reductase للبكتيريا والفطريات على مجموعة غير بروتينية أخرى وهي

الـ (Flavin Adenin Nucleotide) FAD و التي لا توجد في الـ Nitrite reductase النباتي (Siegel et al., 1989).

2.6_نبذة تاريخية عن استخلاص الـ Nitrite-reductase

لقد تم استخلاص إنزيم الـ Nitrite reductase من عدة أنواع نباتية مثل السبانخ Spinach (Vega et al., 1977) و الشعير Barley (Serra et al., 1982) و البازلاء Pea (Bowsher et al., 1988) و القمح Wheat (Small et al., 1984) و نبات اليقطين (Hucklesby et al., 1976) Vegetable Marrow.

كما تم عزل الجينات المشفرة لإنزيم الـ Nitrite reductase للعديد من النباتات الراقية وتم استساخها وتحديد تتابعها، حيث تم عزل الجينات المشفرة له عند نبات السبانخ Spinach (Back et al., 1988)، الذرى Maize (Lahners et al., 1988) و نبات شجر الحكمة (Friemann et al., 1992)، و التبغ Tobacco (Kronenberger et al., 1993) و نبات Arabidopsis thaliana (Haas et al., 2003)، الفاصوليا Bean (Sander et al., 1995)، الرز Rice (Terada et al., 1995)، الفول Soybean والفلفل الأحمر Red Pepper (Kim et al., 1999) واللوتس Lotus (Orea et al., 2001) و الخوخ Peach (Nakamura et al., 2001).

كما تم تحديد تتابع الأحماض الأمينية لـ Nitrite reductase عند كل من:

1- البكتيريا الزرقاء Cyanobacteria:

(Luque et al., 1993) *Synechococcus P.C.C. 7942-*

(kaneko et al., 1996) *Synechocystis sp. P.C.C. 6803-*

(Merchan et al., 1995) *Phormidium Laminosum-*

(Nakamura et al., 2002) *Thermosynechococcus Elongates BP-1-*

2- البكتيريا Bacteria:

(White et al., 1999) *Deinococcus radiodurams-*

(Ivanova et al., 2003) *Bacillus cereus-*

(Gloeckner et al., 2003) *pirellula sp-*

(Fitz-Gibbon et al., 2002) *Pyrobaculum Aerophilum-*

3- الطحالب Algae: *Bigelowiella natans* (Archibald et al., 2003).

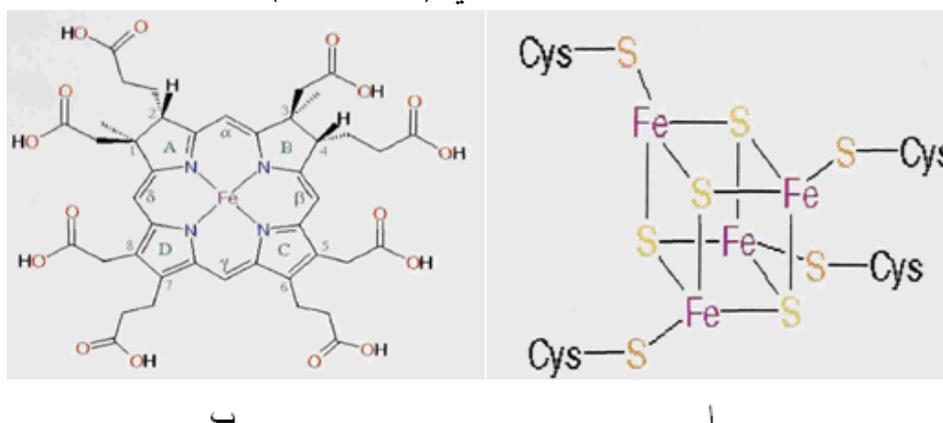
يتمركز الـ Nitrite reductase لأوراق النباتات الراقية في المادة الأساسية للكوروبلاست حيث يتميز ببنية خاصة (Knaff et al., 1996)، حيث تم استنساخ و تحديد التسابع الكامل لأحماضه الأمينية عن طريق تسابع cDNA وتحديد تسابع النهاية الأمينية انطلاقا من البروتين الناضج (Back et al., 1988).

يحتوي الـ Nitrite reductase لنبات السبانخ على 594 حمض أميني تتضمن 32 حمض الأخيرة في النهاية الأمينية ببنتيد العبور، يوضح (الشكل 19) التسابع الكامل للأحماض الأمينية :

MASLPVNKIIPSSSTLLSSSNRRRRNNSIRCQKAVSPAETA AVSPSVDAARLEPRVEERDGFVWLK
EEFRSGINPAEKVKIEKDPMKLFI EDGISDLATLSMEEVDKSKHNKDDIDVRLKWLGLFHRRKHHYGRF
MMRLKLPNGVTTSEQTRYLASVIKKGKDG CADVTTRQNWQIRGVVLPDVPEIIKGLSEVGLTSLQSGM
DNVRNPVGNPLAGIDPHEIVDTRPFTNLISQFVTANSRGNLSITNLPKRWNP CVIGSHDLYEHPHINDL
AYMPATKNGKFGFNLLVGGFFS IKRCEEAIPLDAWVSAEDWPVCK/Uy i LEAFRDLGFRGNRQKCRMMW
LIDELGMEAFRGEVEKRMPEQVLERASSEELVQKWERREYLG VHPQKQQGLSFVGLHIPVGR LQADEME
ELARIADVYGS GELRLTVEQNIIPNVENSKIDSLLEP LLKERY SPEPPI LMKGLVACTGSGFCGQAI
IETKARALKVTEEVQRLVSVTRPVRMHWTGCPNSCGVQVADIGFMGCMTRDENGKPC EGADVFVGGRI
GSDSHLGDLYKKA VPC KDLVPVVAEILINQFGAVPRERE EAE

الشكل (19): يبين التسابع الأميني لـ Nitrite reductase عند نبات السبانخ Spinach. 32 حمض أميني الأولى هي عبارة عن Transit peptide.

الـ Nitrite reductase عند كل من النباتات و الطحالب و البكتيريا الزرقاء هو بروتين ذو سلسلة بيبتيديية واحدة Monomeric وبوزن جزيئي يقارب الـ 63 كيلو دالتون (Back et al., 1991; Hilliard et al, 1988)، كما يحتوي على مجموعتين غير بروتينيتين Prosthetic Groups في موقعه الفعّال Site active، مجموعة واحدة من الـ Siroheme (الشكل 20-ب) (Hirasawa et [Fe₄-S₄]، ومجموعة واحدة من (Vega et al., 1977; Ida et al., 1986) (الشكل 20-أ): (Ip et al., 1990; al., 1987a)

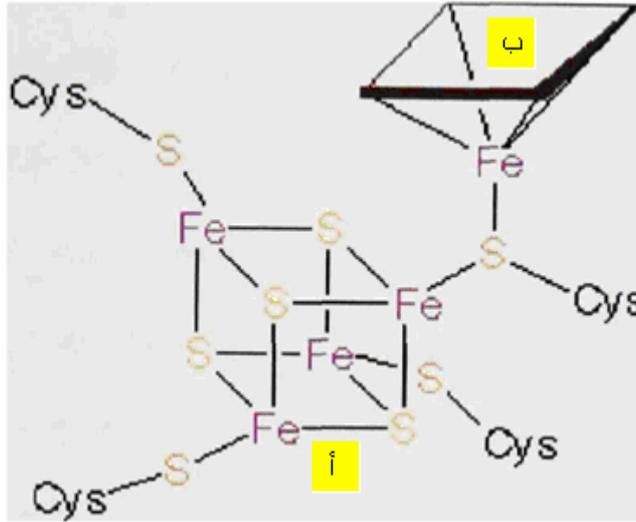


الشكل (20): يبين الصيغة الفراغية لكل من المجموعتين غير الروتينيتين مجموعة [Fe₄-S₄] (أ) ومجموعة الـ Siroheme (ب).

3.6_ مقارنة بين إنزيم الـ Sulfite reductase و Nitrite reductase

إنَّ إنزيم الـ Sulfite reductase الذي يحفز اختزال (SO_3^- إلى SO_2^-) بواسطة ستة إلكترونات يحتوي أيضا على نفس المجموعات غير البروتينية في موقعه الفعّال التي يحتوي عليها إنزيم الـ Nitrite reductase (Siegel et al., 1982; Krueger et al., 1982) و (Janick et al., 1982; Krueger et al., 1982). وقد دل التحليل الطيفي لإنزيم الـ Sulfite reductase عند الـ E.coli على وجود جسر قوي مشترك بين المجموعتين غير البروتينيتين الـ Siroheme و مجموعة $[\text{Fe}_4\text{-S}_4]$.

كما تم إظهار نفس الارتباط بين مجموعة الـ Siroheme ومجموعة $[\text{Fe}_4\text{-S}_4]$ عند الـ Nitrite reductase لنبات السبانخ (Christner et al., al.,1981; Wilkerson et al., 1983) كما هو موضح في (الشكل 21):



الشكل (21): يبين ارتباط المجموعتين غير البروتينيتين مجموعة $[\text{Fe}_4\text{-S}_4]$ (أ) ومجموعة الـ Siroheme (ب) بواسطة حمض أميني مشترك Cys.

كما دلت معطيات بلورة الـ Sulfite reductase المؤكسد عند الـ E.coli بتصميم 3.0 \AA على أن كل من مجموعة الـ Siroheme ومجموعة $[\text{Fe}_4\text{-S}_4]$ قريبتان ومتقابلتان بنيويا داخل الإنزيم، ويبدو بوضوح أن الارتباط يكون بين ذرة الكبريت للسلسلة الجانبية للسيستين Cys التي ترتبط بذرة الحديد المركزية لمجموعة الـ Siroheme وإحدى ذرات الحديد لمجموعة $[\text{Fe}_4\text{-S}_4]$ ، حيث تقدر المسافة بين ذرة الحديد المركزية للـ Siroheme بمركز مجموعة الـ $[\text{Fe}_4\text{-S}_4]$ ب 0.55nm والمسافة بين ذرة الحديد للـ Siroheme وأقرب ذرة حديد لمجموعة $[\text{Fe}_4\text{-S}_4]$ هي 0.44nm (McRee et al., 1986).

كما يوجد موقع ثانٍ للارتباط بين المجموعتين غير البروتينيتين وذلك بين واحدة من ذرات الكبريت لمجموعة $[Fe_4-S_4]$ والتي تصنع رابطة فان دار فالس Van Der Waals مع حافة الحلقة الكبيرة لمجموعة الـ Siroheme (McRee et al. 1986).

هذه هي البنية الوحيدة التي تمكن من تفسير إرجاع الإلكترونات المتعددة لهذا النوع من الإنزيمات، كما أن بنية مجموعة الـ Siroheme التي تدخل في تركيب تحت وحدة الـ Sulfite reductase عند E.Coli كشفت عن وجود شبكة واسعة للروابط الهيدروجينية الناتجة عن تفاعلات جزيئات الماء، ومجموعات الكربوكسيل للسلاسل الجانبية لرباعيات البيروكس الخاصة بمجموعة الـ Siroheme، الشحنة الموجبة للسلاسل الجانبية للأحماض الأمينية وكذا مادة التفاعل السيلفييت (Crane et al., 1997).

ولقد أُقترح أيضاً أنّ هذه الشبكة من الروابط الهيدروجينية تلعب دوراً محورياً في قابلية الإنزيم اختزال الإلكترونات من Sulfite إلى Sulfide (SO_3^- إلى SO_2^-) بواسطة 6 إلكترونات بدون استعمال أي وسائط أثناء دورة التحفيز (Crane et al., 1997).

إنّ المطابقة بين التتابع الأميني لإنزيم الـ Nitrite reductase عند السبانخ وتتابع إنزيم الـ Sulfite reductase عند البكتيريا كشفت أنّ الأحماض الأمينية من نوع سيستين Cys محفوظة كلياً بينهما (Bellissimo et al., 1995).

كما أثبتت دراسات تجارب التطهير المباشر Site-Directed Mutagenesis لإنزيم الـ Nitrite reductase المستنسل أنّ تواجد السيستين Cys في الموقع 514 و 518 في الإنزيم الناضج جد ضرورية لربط مجموعة $[Fe_4-S_4]$ ونجاح نشاط الإنزيم (Bellissimo et al., 1995).

اعتماداً على المقارنات العديدة بين الإنزيمات التي لها نفس العوامل المساعدة فإن شطر تتابع النهاية الكربوكسيلية للـ Nitrite reductase تسمح بالاحتواء على مراكز مؤكسدة. كما افترض أنّ شطر النهاية الأمينية الذي يتضمن عدداً من الأحماض الأمينية ذات الشحنة الموجبة يمكن أن تشارك في ربط المادة المانحة للإلكترونات ذات الشحنة الحامضية (السالبة) الكبيرة وهي الفيرودوكسين Ferredoxin (Campbell et al., 1990; Friemann et al., 1992).

ومع ذلك فإنّ (Dose et al., 1997) عمد باستخدام تجارب التعديل الكيميائي إلى تحديد ثلاث أحماض أمينية ذات شحنة موجبة R375، R556، K436 في إنزيم الـ Nitrite reductase عند السبانخ التي من الممكن أن تشترك في ربط الـ Ferredoxin مما يدل على أنّ هذه المناطق من البروتين هي التي يرتبط بها الـ Ferredoxin.

4.6_ بناء مجموعة الـ Siroheme

إنّ مجموعة ال Siroheme هي عبارة عن ذرة حديد مركزية ترتبط برباعي البيرول Tetrapyrrole والتي لها علاقة من الناحية البنوية بالكلوروفيل Chlorophyll و الكوبالمين Cobalamin، حيث تحتاج مجموعة ال Siroheme إلى الجين CysG لتركيبها (Goldman et al., 1993)، ولأنّ الجين CysG يُشفر لبناء الإنزيم Uroporphyrinogen III methyltransferase [EC 2.1.1.107] فإنّ هذا الأخير هو المفتاح الضروري لتركيب مجموعة ال Siroheme من (Uroporphyrinogen III Warren et al., 1990). وقد استنسل cDNA الكامل المشفر لتركيب ال Uroporphyrinogen III Methyltransferase في بكتيريا E.coli (Peakman et al., 1990).

5.6_ موقع ارتباط مادة التفاعل النتريت NO_2^-

مجموعة كبيرة من الأدلة المتوصل إليها عن طريق الدراسات الطيفية Spectroscopic أثبتت أنّ مادة التفاعل النتريت NO_2^- ترتبط بذرة الحديد المركزية لمجموعة ال Siroheme في إنزيم ال Nitrite reductase المؤكسد محولا بذلك ال Siroheme الحديدي ذو العزم الكبير (High-spin ferric siroheme) الموجود في الإنزيم إلى شكل Siroheme حديدي ذو عزم منخفض (Low-spin ferric siroheme) في الإنزيم (Hirasawa et al., 1987a; Day et al., 1987; Kuznetsova et al., 2004).

ولقد أثبت كل من (Aparicio et al., 1975; Vega et al., 1977; Cammack et al., 1978; Hirasawa et al., 1980; Hirasawa et al., 1987a; Mikami et al., 1989) أنّ مادة التفاعل النتريت NO_2^- أو مواد الارتباط المثبطة التي ترتبط بمجموعة الهيم كالسيانيد Cyanide (Hirasawa et al., 1981)، وأحادي أكسيد الكربون Carbon monoxide (Vega et al., 1977) قد ثبت أنّها تزيل تماما أو تخفض العزم العالي المميز للهيم الحديدي (High-spin ferric heme).

كما أن تجربة طيف الرنين النووي المغناطيسي للإنزيم و مادة الارتباط المثبطة السيانيد Cyanide أضافت دليل آخر على أنّ الارتباط بين مواد الارتباط هذه في الإنزيم يؤدي إلى الانتقال من العزم العالي إلى الضعيف (Ondrias et al., 1985).

وقد أظهر طيف الرنين المغناطيسي للإلكترون EPR الفائق الدقة لـ NO^{15} و NO^{14} المضافة لمجموعة ال Siroheme أنّها ترتبط بذرة الحديد المركزية لها (Fry et al., 1980; Ip et al., 1990).

6.6_آلية انتقال الإلكترونات المرجعة لمادة التفاعل النتريت

إنّ الآلية التي يتم بموجبها نقل الإلكترونات وتحويل النتريت إلى الأمونيا مباشرة بواسطة ناقلي الإلكترونات مجموعة الـ Siroheme و [Fe₄-S₄] بدون تحديد وسائط ثابتة تبقى غير معروفة لحد الآن، أما تجارب التحليل الضوئي بالليزر فقد أثبتت الدليل الأول لانتقال الإلكترونات من مجموعة الـ [Fe₄-S₄] إلى مجموعة الـ Siroheme (1994 Siroheme Hirasawa et al.) وأما الدراسات التي تم فيها استعمال مادة الـ Dithionite كمعطي تجريبي للإلكترونات قد أثبتت أنّ مجموعة [Fe₄-S₄] المختزلة تُعاد أكسدها سريعاً بواسطة النتريت-NO₂⁻ Lancaster et al., 1979).

إنّ ارتباط النتريت-NO₂⁻ بالحالة المؤكسدة للـ Nitrite reductase يكون بطيء نسبياً وهذا ما يدل على إمكانية اختزال مجموعة الـ Siroheme إلى الحالة الحديدية (Ferrous state) قبل ارتباطها بالنتريت (Cammack et al., 1978). هناك دليل آخر يدل على أنّ معقد الـ Siroheme-NO الحديدي يمثل اختزال النتريت-NO₂⁻ بواسطة إلكترونين والذي يمكن اعتباره كوسيط أولي (Kuznetsova et al., 2004; Aparicio et al., 1975; Cammack et al., 1978; Fry et al., 1980).

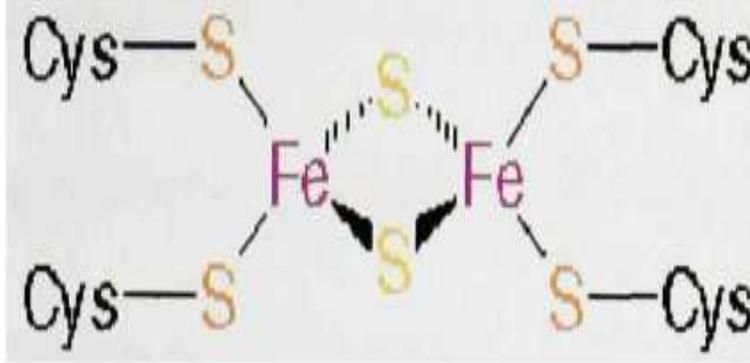
بينما الـ Hydroxylamine (NH₂OH) الذي له إلكترونين مؤكسدين أكثر من الـ Ammonia يعتبر كوسيط متأخر (Kuznetsova et al., 2004; Vega et al., 1977). لقد وجد أيضاً أنّ إنزيم الـ Nitrite reductase المؤكسد يتفاعل مع الـ (NH₂OH) Hydroxylamine لينتج نفس المعقد الـ Siroheme-NO الحديدي المشكّل أثناء حضانة الإنزيم والنتريت-NO₂⁻ مع مُرجع تجريبي (Kuznetsova et al., 2004).

وفي الآونة الأخيرة دلت دراسات الامتصاص الطيفي وطيف الرنين المغناطيسي للإلكترون EPR للدورة التحفيزية عند الـ Nitrite reductase لدى نبات السبانخ Spinach مع إنزيم الفيروكسين Ferredoxin المرجّع بواسطة النظام الضوئي الأول Photosystem I كمعطي الكترولونات، أشارت إلى أنّ نسبة الحد من تفاعل الـ Nitrite reductase و الـ Hydroxylamine ترجع إلى خطوة ارتباط Hydroxylamine (Kuznetsova et al., 2004).

هذه الدراسة اقترحت أيضاً أنّ معقد الـ NiR-hydroxylamine تمثل نقطة تفرع في حلقة التحفيز عند الإنزيم، حيث أنّ في حالة غياب المرجع فإنّ الإنزيم يرجع إلى الحالة Fe⁺⁺ Siroheme-NO أما في وجود المرجع Reductant فإنّه يدخل في تشكّل الأمونيا.

7_ إنزيم الفيروكسين Ferredoxin

هو عبارة عن بروتين منحل أحادي السلسلة ذو وزن جزيئي ضعيف (حوالي 12 kd) حيث يعمل كمعطي طبيعي للالكترونات، وهو يحتوي على مجموعة غير بروتينية واحدة متمثلة في مجموعة الـ $[Fe_2-S_2]$ (Cammack et al., 1984)، حيث أنّ مجموعة الـ $[Fe_2-S_2]$ تتكون من ارتباط ذرتي حديد على شكل هرم رباعي المستويات Coordinated Tetrahedrally مع بعض بواسطة ذرتي كبريت غير عضويتين وكذا روابط مع 4 ذرات كبريت للسلاسل الجانبية لأربع أحماض أمينية سيستين Cys كما هو موضح في (الشكل 22):



الشكل (22): بنية مجموعة $[Fe_2-S_2]$ حيث تتكون من ذرتي حديد مرتبطة عن طريق جسرين كبريتيين من جهة ومن جهة أخرى بأربع ذرات كبريت للأحماض الأمينية سيستين Cys .

أثناء مرحلة التركيب الضوئي يُرجع الفيريدوكسين Ferredoxin بواسطة النظام الضوئي الأول PhotosystemI فيصبح فيريدوكسين مُرجع حيث يعمل بعد عملية الإرجاع هذه كمانح للالكترونات للعديد من الإنزيمات المتدخلة في عمليات البناء الحيوي كتمثيل الأروت وتمثيل الكبريت و إزالة تشعب الأحماض الدهنية وكذا بناء الغلايسين بيتاين ... الخ (Arnon et al., 1988).

يوجد العديد من الإنزيمات التي تم عزلها من النباتات و الطحالب و البكتيريا الزرقاء التي تستعمل الفيريدوكسين المرجع كمانح للالكترونات، ومن بين هذه الإنزيمات نجد (Nir) Ferredoxin:nitrite reductase، وكذا Ferredoxin:sulfite reductase (SiR)، Ferredoxin:thioredoxin reductase (FTR)، Ferredoxin:thioredoxin reductase (FNR)، Ferredoxin: NADP Glutamate synthase (GOGAT).

كما وجد أنّ الفيريدوكسين Ferredoxin يعمل أيضا كمانح للالكترونات عند البكتيريا الزرقاء وذلك قصد اختزال النترات إلى نترات (Schmitz et al., 1995).

كما وجد أنّ الفيريدوكسين المتواجد في الأنسجة غير اليخضورية كخلايا الجذور والفواكه مميز شكلا عن الفيريدوكسين الموجود في الأوراق (Morigasaki et al., 1990).

توجد عدّة أنواع من الفيريدوكسين والتي تختلف عن بعضها البعض في عدد ونوع المجموعات غير البروتينيّة [Fe₂-S₂] فمنها ما يحتوي على مجموعة واحدة ومنها ما يحتوي على مجموعتين، كما أنّ هناك ما يحتوي على المجموعة [Fe₃-S₄] ومنها ما يحتوي على المجموعة [Fe₄-S₄] أو كلاهما معا (Dai et al., 2000; Meyer et al., 2001).

يحتوي النّبات على نوع من الفيريدوكسين Ferredoxin ذو المجموعة غير البروتينيّة الواحدة وهي الـ [Fe₂-S₂] (Cammack et al., 1985) حيث أنّ مجموعة الـ [Fe₂-S₂] ترتبط بالبروتين بأربع أحماض أمينية سيستين Cys.

يتواجد فيريدوكسين النّبات في المادّة الأساسيّة للصّناعة الخضراء حيث تُرجع المجموعة [Fe₂-S₂] بواسطة الالكترونات العابرة لغشاء التيلاكويد من البلاستوسيانين نحو النّظام الضّوئي الأوّل (Knaff et al., 1991).

8_ آليّة ارتباط الـ Ferredoxin بالـ Nitrite reductase

يتشكّل معقّد الكتروستاتيكي بين معطي الالكترونات الفيريدوكسين والإنزيمات الهدف بصورة عامّة، حيث وُجد أنّ تشكّل هذا المعقد يكون بين الفيريدوكسين و FNR (Foust et al., 1969) وبين الفيريدوكسين و GOGAT (Knaff et al., 1991) وبين الفيريدوكسين و FTR (Hirasawa et al., 1988) وكذا بين الفيريدوكسين والـ Nitrite reductase (Schmitz et al., 1995) وحديثاً تمّ تأكيده بشكل واضح بين الفيريدوكسين والـ Nitrate reductase (Hirasawa et al., 2004).

وقد بيّنت الأدلّة على أنّ هناك تشابه في آليّة الارتباط بين الفيريدوكسين والـ Nitrite reductase وبين الفيريدوكسين و GOGAT وبين الفيريدوكسين و FNR (Hirasawa et al., 1986, 1989).

بالنسبة للـ Nitrite reductase فإنّه يشكّل معقّد الكتروستاتيكي ثابت مع الفيريدوكسين حيث يساهم الفيريدوكسين من جهة بمجموعة من الأحماض الأمينية السّالبة الشّحنة بينما يساهم الـ Nitrite reductase بمجموعة مكملة للمعقد من الأحماض الأمينية ذات الشّحنة الموجبة من جهة أخرى (Knaff et al., 1996; Hirasawa et al., 1985).

الطاقة

1_ النّمدجة بالتّمائل

1.1_ جهاز الإعلام الآلي

للقيام بعملية النمذجة بالتماثل Homology modeling نحتاج إلى جهاز حاسوب يكون ذو مواصفات جيّدة، حيث يجب أن يتمتع بذاكرة تساوي أو تفوق 512Mb ومساحة تفوق Mb 100 على القرص الصلب وهي مساحة حرة تكفي لعمل البرامج المختلفة المستعملة في عملية النمذجة بالتماثل، وكذا البرامج المستعملة في عرض البنيات الثلاثية أو التي ستستخدم في تقييم النماذج البروتينية التي نحصل عليها، وأيضا الملفات الخاصة بالخوارزميات والبنيات البروتينية. يعمل على نظام تشغيل Microsoft Windows XP 2003 أو Linux/ Unix . كما يجب أن يكون موصولاً بالانترنت، وهذا لتنزيل العديد من الملفات والبرامج الالكترونية، وتشغيل بعض البرامج .

2.1_ بنوك المعطيات

1.2.1_ بنك معطيات البروتينات PDB «Protein Data Bank»

وهو عبارة عن بنك معطيات يحتوي على معلومات عن كل جزيئات البروتينات التي تم تحديد بنيتها ثلاثية الأبعاد تجريبياً، بواسطة الأشعة X أو الرنين النووي المغناطيسي، وموقعه على شبكة الانترنت هو <http://www.rcsb.org/pdb/home/home.do>. أين يتم تحميل البنيات ثلاثية الأبعاد للبروتينات القالب وكل المعلومات المفصلة الخاصة بها.

2.2.1_ بنك NCBI "National Center for Biotechnology Information"

يعتبر من أهم قواعد المعطيات البيولوجية المتعلقة بالتتابعات الجينية و التتابعات البروتينية، موقعه على شبكة الانترنت هو <http://www.ncbi.nlm.nih.gov>. أين يتم تحميل تتابع البروتين الهدف الذي نريد بناء بنيته ثلاثية الأبعاد.

3.1_ البرامج المستعملة

تحديد البنية ثلاثية الأبعاد للبروتينات يبقى محدوداً وذلك لتوفر الطرق التجريبية (الرنين النووي المغناطيسي و الأشعة X) في عدد قليل ومحدود فقط من المخابر، و كذا لصعوبة وطول المدة

التي يتطلبها تحديدها تجريبيا، ولهذا اتجه الباحثون للبحث عن وسائل بديلة تعوض هذا النقص و تحاكي التجارب المخبرية وذلك ببناء نماذج نظرية التي تعتمد أساسا على الطرق الحاسوبية والتي نجد من بينها طريقة النمذجة بالتمائل (Baker et al., Homology modeling) (2001).

تعتمد طريقة النمذجة بالتمائل على عدة برامج الحاسوبية، منها برنامج Modeller وذلك لبناء نماذج نظرية ثلاثية الأبعاد للبروتينات التي تم تحديد تتابعاتها الأمينية ولم تحدد لها البنية ثلاثية الأبعاد تجريبيا (Fiser et al., 2000).

يتم تزويد أو إدخال التتابع الأميني للبروتين الهدف الذي نريد بناء بنيته ثلاثية الأبعاد، وكذا بنية بروتين أو أكثر من البروتينات التي نستعملها كقالب ، وكذا عدة ملفات تحمل معلومات خاصة بالبروتين القالب والهدف.

1.3.1_برنامج Modeller

هو برنامج حاسوبي يستعمل للتنبؤ و بناء نماذج نظرية ثلاثية الأبعاد للبروتينات غير محددة البنية تجريبيا، حيث يتم تحضير الملفات التي يتم إدخالها لبرنامج الـ Modeller كتتابع البروتين الهدف، وملفات خاصة ببنية القالب والإحداثيات الذرية لها. يبني الـ Modeller آليا نموذجا واحدا أو عدة نماذج ثلاثية الأبعاد توافق الخصائص الكيميائية الفراغية (طول الروابط، الزوايا و تداخل الذرات ,,, الخ) انطلاقا من التشابه بين التتابع الأميني للبروتين الهدف والبروتين القالب (S´anchez et al., 1997).

إن استعمال هذا البرنامج يمكن من تحديد بنية افتراضية نظرية لتتابع بروتين معين غير محدد البنية ثلاثية الأبعاد انطلاقا من نموذج بروتين آخر محدد التتابع والبنية ثلاثية الأبعاد بعد سلسلة من عمليات المقارنة والحسابات (Sali et al., 1993; Fiser et al., 2003; Sali et al., 1993; John et al., 2003).

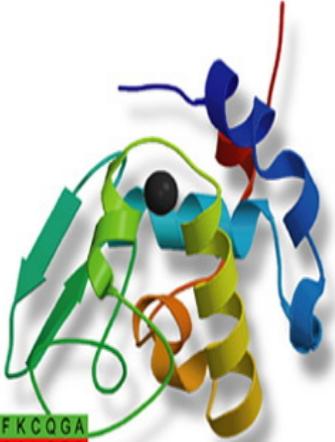
يتم تحميل برنامج Modeller 9v5 بعد الاشتراك من شبكة الانترنت مباشرة من الموقع <http://salilab.org/modeller>، ثم يثبت بعد ذلك على الحاسوب، كما يتم تحميل عدة ملفات خاصة بعمل الـ Modeller من نفس الموقع.

- يستعمل برنامج Modeller لغة الـ Python كلغة لقراءة البرامج الملحقة (Plug in).

- يعمل برنامج الـ Modeller بواسطة كتابة الأوامر مباشرة على واجهة الـ DOS (شكل-2) مستعملا ملفات السندات والبنىات الخاصة بالبروتين الهدف والقوالب التي تم تركيبها.

Modeller

Program for Comparative Protein
Structure Modelling by Satisfaction
of Spatial Restraints



```
A I L V G S M P R R D G M E R K D L L K A N V K I F K C Q G A  
V E V C P V D C F Y E G P N F L V I H P D E C I D C A L C E P  
G A C K P E C P V N I I Q G S . . Y A I D A D S C I D C G S  
C . . A C G A C K P E C P V N I I Q G S . . I Y A I D A D S
```

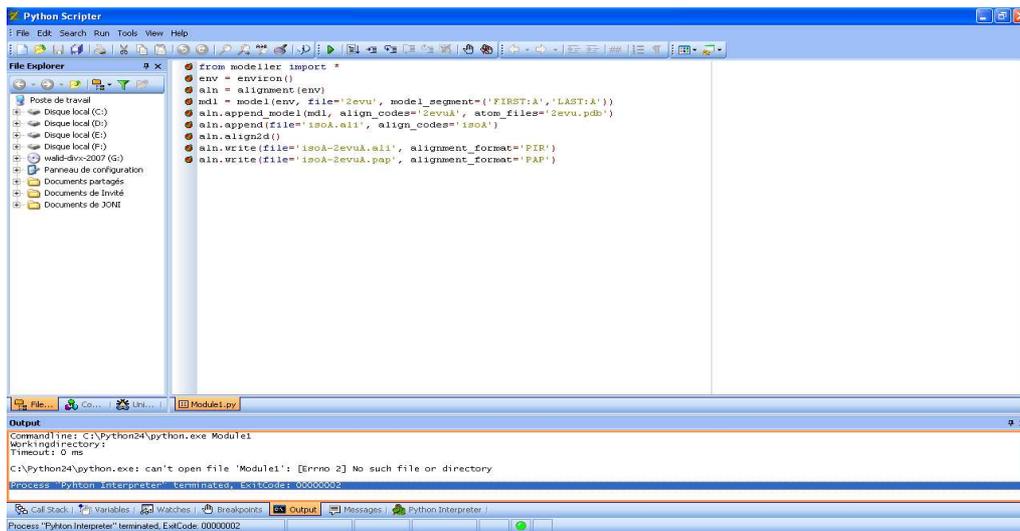
شكل(1): صورة لبرنامج Modeller

2.3.1_برنامج Python

يستعمل برنامج الـ Modeller خوارزميات ديناميكية ذاتية (et Smith et al., 1981; Eswar al., 2005 نقوم ببرمجتها حسب خصائص البروتين الهدف وجزئيات البروتين المستعملة كقالب، نقوم في الغالب ببرمجتها بواسطة لغة الـ Python هذا الأخير يتم تحميله من الموقع <http://www.python.org/download/>



شكل (2): نافذة العرض DOS لبرنامج Modeller 9v5.

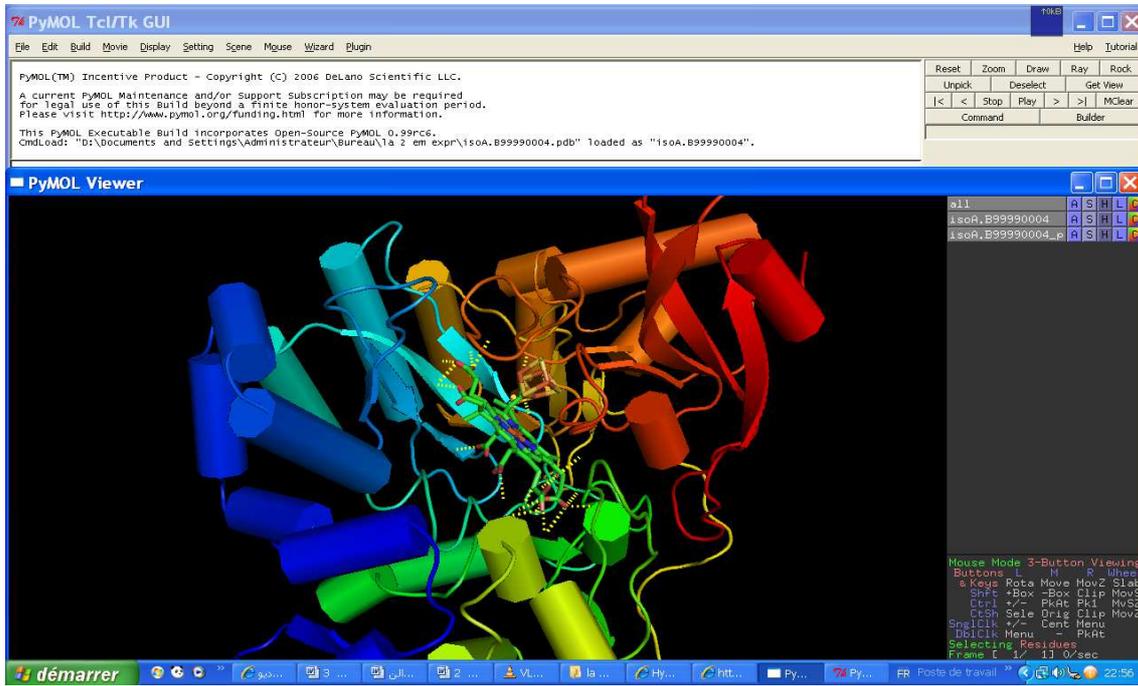


شكل (3): النافذة الأساسية لبرنامج Python Scripter المستعمل في برمجة Script التي يتم تنفيذها بواسطة برنامج Modeller.

3.3.1 برنامج Pymol

يستعمل برنامج Pymol لتصميم وعرض البنيات ثلاثية الأبعاد للبروتينات (Rafael et al., 2008)، وهو متوفر بعدة نسخ منها الشكل المجاني على الموقع الخاص www.pymol.com، حيث يمكن تحميله ونقله من جهاز حاسوب إلى آخر، وتشغيله على عدة أنظمة حاسوبية .MacOS .windows xp. Windows vista .linux.

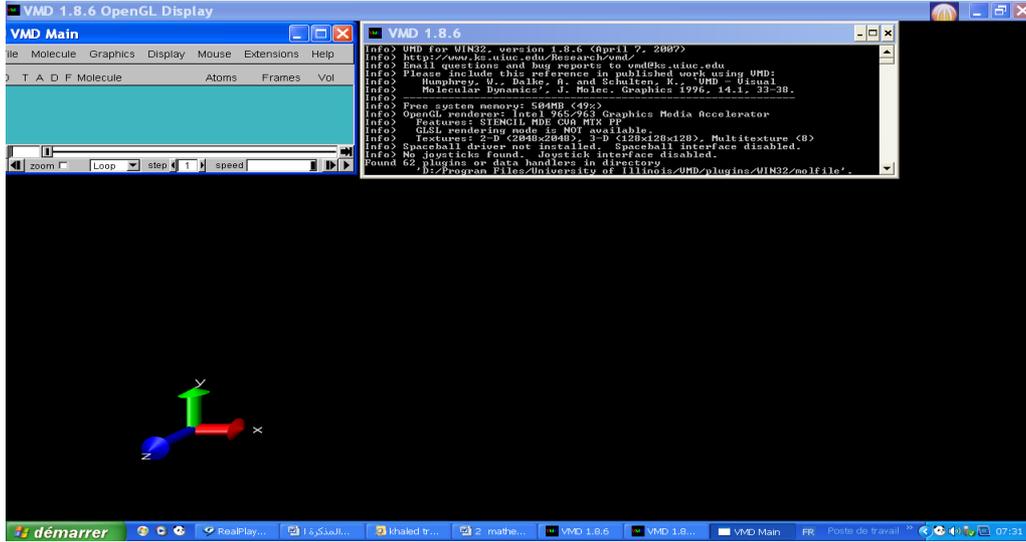
يتميز برنامج الـ Pymol عن غيره من برامج العرض بعدة خصائص تمكنه من معالجة البنيات الثلاثية الأبعاد بشكل جيد ومتعدد كتوضيح توضعها الفراغي، و تحليل بعض الخصائص الكيميائية لها.



شكل(4): نافذة العرض لبرنامج Pymol.

4.3.1 برنامج VMD

يعتبر من بين أهم البرامج في عملية عرض الجزيئات وتحليل بنيات البروتين كما يستعمل لتقييم النموذج البروتيني المتحصل عليه وذلك بقياس RMSD وكذا العامل Q (Humphrey et al., 1996).



شكل(5): نافذة العرض لبرنامج VMD

2_ مراحل عملية النمذجة بالتمثيل

1.2_ الخطوة الأولى: الحصول على تتابع البروتين الهدف وتحديد القالب المناسب

يتم الحصول على التتابع الأميني للبروتين الهدف Target، إما بتحديد التتابع تجريبيا بعد استخلاص البروتين، أو تنزيله بشكل ملف إلكتروني بامتداد "Blast" (Altschul et al., 1997) مباشرة من بنك المعطيات NCBI. بعد ذلك نقوم بالبحث عن تتابعات كل البروتينات (محددة البنية تجريبيا) والتي لها نسبة تشابه معتبرة بتتابع البروتين الهدف، نختار التتابعات التي لها نسبة تشابه كبيرة مع تتابع البروتين الهدف Target والذي نستعمله كقالب Template. تعتبر هذه الخطوة سهلة نتيجة توفر عدد كبير من قواعد المعطيات للتتابعات الأمينية وكذا برامج البحث اللازمة لذلك (Holm et al., 1994; Barton et al., 1998; Altschul et al., 1996), نتحصل على التتابعات الأمينية للبروتينات القوالب الملائمة بإدخال تتابع البروتين الهدف كموضوع بحث في إحدى قواعد المعطيات التي تحوي البنيات ثلاثية الأبعاد مثل بنك معطيات البروتينات PDB (Abola et al., 1987 ; Berman et al., 2000) أو SCOP (Hubbard et al., 1999) أو DALI (Holm et al., 1999) أو CATH (Orengo et al., 1999) كما هو موضح في (الجدول 1).

الجدول (1): بعض مواقع قواعد البيانات على شبكة الانترنت الخاصة بالبنيات البروتينية.

الموقع على شبكة الانترنت	قاعدة البيانات
www.biochem.ucl.ac.uk/bsm/cath	CATH
www.ncbi.nlm.nih.gov/GenBank	GenBank
bioinfo.mbb.yale.edu/genome	GeneCensus
guitar.rockefeller.edu/modbase	MODBASE
www.rcsb.org/pdb/	PDB
presage.stanford.edu	PRESAGE
scop.mrc-lmb.cam.ac.uk/scop/	SCOP
www.ebi.ac.uk/swissprot	SWISSPROT+TrEMBL

هناك ثلاثة أقسام رئيسية من طرق مقارنة البروتينات لتحديد البروتين القالب المناسب:

المجموعة الأولى من هذه الطرق تتضمن مقارنة تتابع البروتين الهدف بتتابع كل بروتين من بروتينات قواعد البيانات على حدى وذلك باستعمال مقارنة الثنائيات تتابع-تتابع Pairwise sequence-sequence (Apostolico et al., 1998). تكمن فعالية هذه الطرق في البحث بشكل مباشر عن تتابعات البروتينات الموافقة (Pearson et al., 1995) والحصول على بنيات ذات مواصفات جيدة (Brenner et al., 1998)، والبرامج الحاسوبية الأكثر استعمالا في هذا المجال هي FASTA (Pearson et al., 2000) والبرنامج BLAST (Altschul et al., 1990).

المجموعة الثانية تعتمد على مقارنة تتابع البروتين الهدف بعدد كبير من تتابعات البروتينات Multiple sequence comparisons وذلك للرفع من دقة البحث و تزيد بشكل واسع من مجموعة النماذج الشبيهة بالتتابع الهدف وتحديد نسبة التشابه بالترتيب حسب نسبة التشابه (Altschul et al., 1997; Gribskov et al., 1996; Henikoff et al., 2000)، يعتبر البرنامج الحاسوبي PSI-BLAST الأكثر استعمالا في هذا المجال (Altschul et al., 1997).

المجموعة الثالثة من هذه الطرق والمعروفة بالـ Threading (Bowie et al., 1991)، تتضمن مقارنة التتابع الأميني للبروتين الهدف ببنية البروتينات المحددة تجريبيا مباشرة بواسطة البرامج الحاسوبية الممثل بعضها في (الجدول 2).

الموقع على شبكة الانترنت	برنامج البحث عن القالب
--------------------------	------------------------

www.lmmb.ncifcrf.gov/_nicka/123D.html	123D
www.ncbi.nlm.nih.gov/BLAST/	BLAST
www2.ebi.ac.uk/fasta3	DALI
www2.ebi.ac.uk/dali/	FastA
bioinformatics.burnham-inst.org	MATCHMAKER
www.embl-heidelberg.de/predictprotein/	PHD, TOPITS
www.came.sbg.ac.at	PROFIT P 57
globin.bio.warwick.ac.uk/_jones/threader.html	THREADER
www.doembi.ucla.edu/people/frsvr/frsvr.html	UCLA-DOE FRSVR

الجدول (2): بعض البرامج الحاسوبية المستعملة في البحث عن القوالب الشبيهة بالبروتين الهدف.

على الرغم من نسبة التشابه العالية بين تتابع البروتين الهدف وتتابع القالب هي الخطوة الأساسية والتي ينتج عنها الحصول على قالب جيد، هناك بعض الجوانب التي يجب أن نأخذها بعين الاعتبار عند اختيار القالب وهي كالآتي:

- دراسة العائلة التي ينتمي إليها كل من البروتين الهدف والقالب، و بناء شجرة التقسيم لها Phylogenetic Tree وكذا إجراء مطابقات متعددة مما يساعد على اختيار بنية القالب الأقرب إلى البروتين الهدف (Felsenstein et al., 1985).
- مقارنة المحيط الذي يتواجد فيه البروتين القالب بمحيط البروتين الهدف، وهذا يعني بشكل أساسي: درجة الحموضة PH، درجة الذوبان و مواد الارتباط ..الخ. لأن التشابه بينهما في هذه الخصائص يرفع من قيمة و مصداقية النموذج الذي نريد بناءه (Marti-Renom et al., 2000).
- بنية القالب المحددة تجريبيا لها أهمية كبيرة في الحصول على نموذج جيد وذلك حسب دقة التصميم Resolution (Marti-Renom et al., 2000).

باستخدام الملفات المدمجة (pdb_95.pir (build_profile.Py والأمر Build_profile يقوم برنامج modeller بتحديد التتابعات الأمينية للبروتينات المشابهة للبروتين الهدف وإجراء عملية المطابقة بينها وترتيبها حسب درجة التشابه (Eswar et al., 2005). مع وضع تتابع البروتين الهدف في ملف من الشكل isoA.ali حتى يتمكن برنامج الـ Modeller من قراءته (Pearson et al., 1990).

```
>P1;isoA
sequence:isoA::::::::::0.00:0.00
PRVELKDGFFILKEKFRKGINPQEKVKIEREPMKLFMENGIEELAKKSMEELDSEKSSKDDIDVRLKWL
GLFHRRKHQYKGFMMRLKLPNGVTTSAQTRYLASVIRKYGEDGCADVTRQNWQIRGVVLPDVPEILKG
LASVGLTSLQSGMDNVRNPVGNPIAGIDPEEIVDTRPYTNLLSQFITANSQGNPDFTNLPRKWNVCVVG
THDLYEHPHINDLAYMPANKDGRFGFNLLVGGFFSPKRCEEAIPLDAWVPADDVLPCKAVLEAYRDLG
TRGNRQKTRMMWLIDELGVEGFRTEVEKRMKPNGLERGSSEDLVKNQWERRDYFGVNPQKQEGLSFVGL
HVPVGRQLQADDMDLARLADTYGSGELRLTVEQNIIPNVETSKTEALLQEPFLKNRFSPEPSILMKGL
VACTGSQFCGQAI IETKLRALKVTEEVERLVSVPRP IRMHWTGCPNTCGQVQVADIGFMGCLTRGEEGK
PVEGADVYVGGRIGSDSHIGEIYKKG.....*
```

الشكل (6): وضع التتابع الأميني الخاص بالـ Ferredoxin- nitrite reductase في ملف من الشكل isoA.ali

كما قمنا بالبحث عن البروتينات المحددة البنية تجريبيا الشبيهة بالبروتين الهدف آليا في قواعد البيانات. وذلك بإدخال التتابع الأميني للـ Ferredoxin- nitrite reductase في برنامج SSEARCH-PDB (Smith et al., 1981). وقاعدة البيانات SWISSMODEL الذي يستخدم عدة برامج حاسوبية كالـ BLAST و الـ HHSEARCH في البحث عن البروتينات الشبيهة.

2.2_ الخطوة الثانية: مطابقة تتابع البروتين الهدف بالبروتين القالب

يطابق في هذه الخطوة تتابع البروتين الهدف ببنية البروتين القالب باستعمال عدة برامج (Baxevanis et al., 1998; Briffeuil et al 1998; Taylor et al., 1996)، إذا كانت نسبة التتابعات المتلائمة بشكل عام تفوق 40% فإنّ المطابقة تكون غالبا صحيحة وناجحة، بينما تكون المطابقة ضعيفة إذا كانت نسبة التتابعات المتلائمة بشكل عام أقل من 40% (Saqi et al., 1999)، وتصبح المطابقة صعبة إذا كانت نسبة التتابعات المتلائمة أقل من 30% (Rost et al., 1999). فكلما تناقصت نسبة التشابه تقل نسبة المطابقة ويزداد عدد الثغرات (Gaps) وعدد الأخطاء في بنية النموذج المتحصل عليه. فمثلا 20% فقط من الأحماض الأمينية يمكن أن تكون متلائمة المطابقة عندما تكون نسبة التشابه بين تتابع الهدف وتتابع القالب 30% (Johnson et al., 1993).

البرنامج الحاسوبي الأكثر شيوعا واستعمالا في مجال مطابقة التتابعات الأمينية هو CLUSTAL (Jeanmougin et al., 1998)، كما يوجد العديد من البرامج المماثلة على شبكة الانترنت الخاصة بهذا المجال الممثل بعضها في (الجدول 3):

الجدول (3): يمثل بعض البرامج الحاسوبية المستعملة في المطابقة بين تتابعات البروتينات.

الموقع على شبكة الانترنت	برنامج المطابقة
dot.imgen.bcm.tmc.edu:9331/ 170	BCM SERVER
www.ncbi.nlm.nih.gov/BLAST	BLAST
blocks.fhrc.org/blocks/blockmkr/	BLOCK MAKER
www2.ebi.ac.uk/clustalw/	CLUSTAL
www2.ebi.ac.uk/fasta3/ 127	FASTA3
pbil.ibcp.fr/	MULTALIN

مع التذكير أنه في الحالات التي تصعب فيها المطابقة من المفيد غالبا استعمال واعتماد عدة بنيات بروتينية كقوالب بدلا من قالب واحد، وكذا معرفة أكبر قدر من المعلومات الخاصة بالتتابع (Barton et al., 1987; Taylor et al., 1994).

قمنا بمطابقة تتابع البروتين الهدف isoA مع بنية البروتين الذي تم اختياره كقالب 2akz بواسطة برنامج Modeller وذلك باستخدام الملف المدمج Align.py والأمر Needleman Align (et al., 1970). المطابقة عن طريق برنامج Modeller تختلف عن طرق المطابقة الأخرى لأنها تأخذ بعين الاعتبار المعلومات الخاصة بالبنية أثناء المطابقة، ويتحقق ذلك من خلال الحد من تغيرات الثغرات gaps حيث تعمل على إزاحة وضع الثغرات إلى الأماكن المفتوحة من البروتين والمناطق المنحنية (مناطق الانعطاف وخارج قطع البنية الثانوية) (Levitt et al., 1992). ونتيجة لذلك يتم التقليل من الوقوع في الأخطاء إلى مايقارب النصف مقارنة بطرق وتقنيات المطابقة الأخرى، ولذا تصبح أكثر أهمية من حيث زيادة نسبة التشابه بين تتابع البروتين الهدف والقالب والتقليل من عدد الثغرات.

ولتأكيد العمل السابق قمنا أيضا بعملية المطابقة بين تتابع البروتين الهدف والبروتين القالب 2akz بواسطة برنامج حاسوبي آخر وهو CLUSTALW (Thompson et al., 1994).

3.2_ الخطوة الثالثة: بناء النموذج

توجد عدة طرق لبناء نموذج نظري ثلاثي الأبعاد للبروتين الهدف:

الطريقة الأولى هي بناء النموذج انطلاقا من عدد قليل من المناطق المتشابهة المتحصل عليها أثناء المطابقة وهذه الطريقة تسمى طريقة النمذجة بواسطة تجميع الجسم الصلب Modeling by Rigid-Body Assembly (Blundell et al., 1987; Greer et al., 1990).

الطريقة الثانية والتي تسمى Modeling by Segment Matching تعتمد في النمذجة على الأجزاء أو القطعة المتلائمة التي تتعلق أساسا بالتقارب بين مواضع الذرات المحفوظة بين القالب و الهدف (Claessens et al., 1989 ; Jones et al 1986).

الطريقة الثالثة تعتمد على إزالة العوائق الفراغية Satisfaction of Spatial Restraints حيث تستعمل المسافات الجيومترية و التقنيات المحسنة للتغلب على العوائق الفراغية للبنية ثلاثية الأبعاد المتحصل عليها عند المطابقة (Brocklehurst et al., 1993). دقة بناء النموذج متماثلة عند جميع هذه الطرق إذا استعملت بشكل جيد، كما أنّ هناك عوامل أخرى كاختيار القالب واتساق المطابقة التي يكون لها عادة الأثر الأكبر في بناء نموذج دقيق خصوصا إذا كانت نسبة التشابه بين تتابع البروتين الهدف والبروتين القالب أقل من 40%، ومع ذلك فإنّ طريقة النمذجة بالتمائل تسمح بقدر من المرونة والأوتوماتيكية ما يجعلها الأحسن والأسرع للحصول على نماذج جيدة، فعلى سبيل المثال تسمح هذه الطريقة بسهولة إعادة الحسابات الخاصة بالنموذج عندما نغير في المطابقة.

طريقة النمذجة بالتمائل تسمح ببناء نموذج نظري ثلاثي الأبعاد باستعمال البرامج الحاسوبية والتي من بينها برنامج Modeller، الذي يعتمد أساسا على طريقة إزالة العوائق الفراغية الممكنة Satisfaction of Spatial Restraints، وذلك بتوحيد الخصائص الفراغية بقواعد عامة نموذجية خاصة ببنية البروتين كطول الروابط ودرجة دوران الزوايا بين بنية البروتين القالب والهدف.

تم بناء النموذج النظري الثلاثي الأبعاد للبروتين الهدف بواسطة برنامج الـ Modeller، وذلك بتزويد هذا الأخير بالملف المدمج الخاص ببناء النموذج النظري "model_single.py" بعد تعديل بياناته بما يوافق خصائص البروتين الهدف والبروتين القالب وعدد النماذج المراد بناؤها (اخترنا بناء 10 نماذج)، ثم تحديد المجموعات غير البروتينية في البروتين القالب وهما مجموعة Siroheme و مجموعة [Fe₄S₄] لتحويلها للبروتين النموذج بواسطة الملف المدمج "read_ligand.py".

والمخطط التالي يوضح الخطوات الأساسية المتبعة في بناء النموذج النظري باستعمال برنامج Modeller عن طريق ازالة العوائق الفراغية Satisfaction of Spatial Restraints .

1. Align sequence with structures

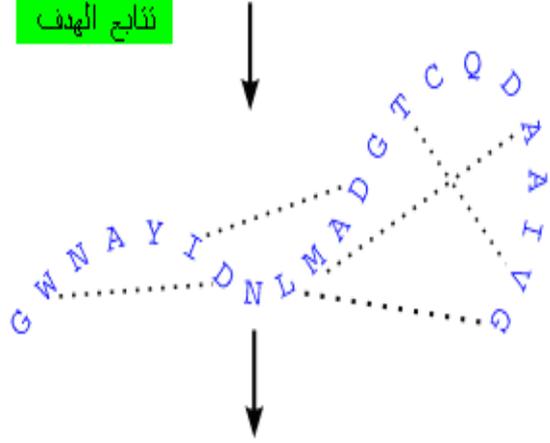
مطابقة نتائج البروتين الهدف
بالبروتين القالب

بنية القالب
Template structure(s)
Target sequence
نتائج الهدف

SWQTYVDTNLVGTGAVTQA--AI
-GWNAYIDNLMADGTCQDAIIVG

2. Extract spatial restraints

ازالة العوائق الفراغية



3. Satisfy spatial restraints

النموذج النظري بعد ازالة العوائق
الفراغية



الشكل (7): مخطط يبين طريقة النمذجة بإزالة العوائق الفراغية بواسطة برنامج Modeller.

4.2_ الخطوة الرابعة: تقييم النموذج المتحصل عليه

بعد بناء النموذج النظري لابد من تقييمه وتحديد الأخطاء الممكنة، لأنه من الممكن جدا أن يحمل العديد من الأخطاء، توجد بشكل عام طريقتين للتقييم يجب إنجازهما (Lüthy et al., 1993; Sippl et al., 1992) وهما:

1.4.2_ التقييم الداخلي Internal evaluation

وهو تقييم التماسك الذاتي Self-Consistency للنموذج المتحصل عليه والذي بموجبه يتم الكشف عن مدى إزالة العوائق الفراغية المعتمدة في بناء النموذج بواسطة برنامج Modeller. يقيم النموذج المتحصل عليه بالاعتماد على خصائص الكيمياء الفراغية للبروتين Stereochemistry (الروابط، درجة الزوايا، دوران الروابط و الروابط غير التكافئية...الخ) وذلك بواسطة عدة برامج حاسوبية أهمها برنامج PROCHECK (Laskowski et al., 1993).

1.1.4.2 _ برنامج PROCHEK

يمكن تقييم خصائص الكيمياء الفراغية (Stereochemistry) لبنية البروتين النموذج بواسطة برنامج PROCHECK، الذي يعتمد على معايير الكيمياء الفراغية المستمدة من التصاميم عالية الدقة للبنيات البروتينية المحددة تجريبيا، بحيث أنّ بنية البروتينات التي سيقارن بها النموذج تكون على أساس حمض أميني بحمض أميني (Morris et al., 1992) ، يشمل هذا البرنامج دراسة مخطط Ramachandran Plot، طول الروابط و درجة الزوايا Bond length and angles، الزوايا Dihedral angles، توافق الذرات Atom-Atom overlaps، الخ (Hooft et al., 1996).

2.4.2_ التقييم الخارجي Evaluation external

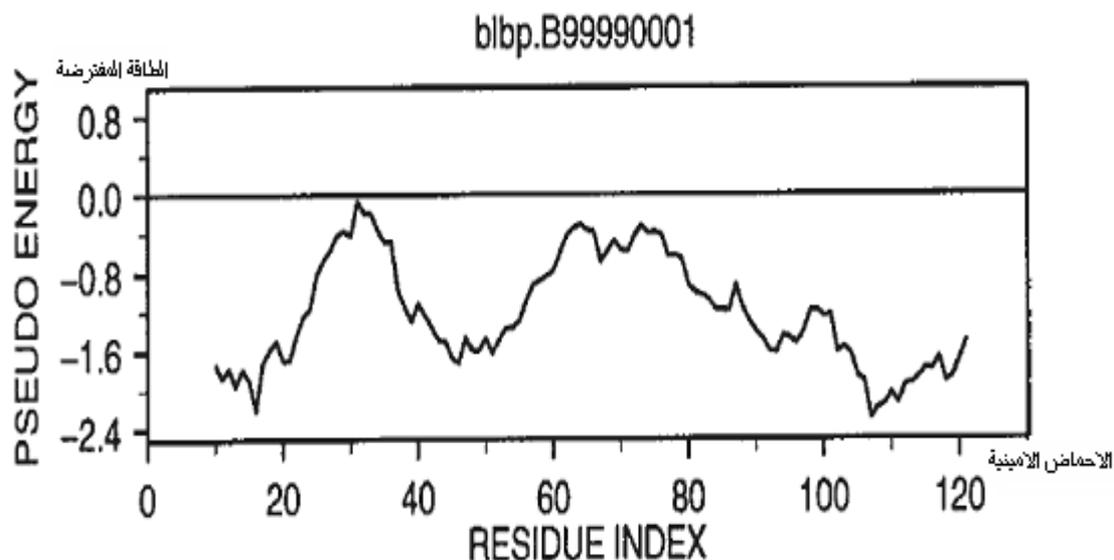
وهو تقييم المعلومات الخاصة بالبنية البروتينية والتي لم يتم اعتمادها أثناء بناء النموذج البروتيني من طرف برنامج Modeller.

إذا كانت نسبة التشابه بين تتابع البروتين الهدف والبروتين القالب المتحصل عليه أقل أو تقارب الـ 30% فإن الغاية الأولى من التقييم الخارجي هي اختبار مدى صحة بنية القالب أو القوالب المستعملة وهي ذات أهمية خاصة إذا كانت نسبة المطابقة هامشية أو في حالة استعمال عدة بنيات مختلفة كقوالب، نسبة التشابه الضعيفة ينتج عنها عادة نماذج تحمل الكثير من الأخطاء مع وجود صعوبة في تمييز ما إن كانت الأخطاء موجودة في البنيات المستعملة كقوالب من جهة أو موجودة في عملية المطابقة من جهة أخرى. و لهذا الغرض توجد طريقة لاكتشاف وجود أو خلو النموذج النظري من الأخطاء وهي بحساب قيم Z-score بواسطة برنامج PROSAII لبنية البروتين النظري المتحصل عليه وبنية البروتين أو البروتينات المعتمدة كقوالب ثم المقارنة بينهما (Sippl et al., 1993). قيمة Z-score هي مقياس لمدى الملائمة بين التتابع الأميني و البنية ثلاثية الأبعاد للبروتين.

1.2.4.2 برنامج PROSAII

يقوم هذا البرنامج بحساب قيمة Z scores لكل من النموذج النظري المتحصل عليه و القالب، هذا من جهة و يقارن قيم Z scores لبنية نموذج البروتين الهدف المتحصل عليها بقيم Z scores بنية البروتين القالب من جهة ثانية. علماً أنّ قيمة Z score للبنية البروتينية هو قياس مدى الملائمة بين تتابع أحماضه الأمينية وبنيته ثلاثية الأبعاد، وعلى الرغم من ذلك فإنّ هذا النوع من التقييم لا يصلح في بعض الحالات حيث يمكن أن يكون لبنية نموذج جيدة مقدار Z-score غير مناسب وهذا لأنّ برنامج PROSAII غير مناسب لبعض الحالات من الانطواء (Sippl et al., 1993).

يقوم برنامج PROSAII ايضاً بكشف المناطق غير الصحيحة في بنية البروتين وذلك برسم المنحنى الطاقوي له Energy profile ، يمثل المنحنى الحالة الطاقوية لكل حمض أميني في البنية أين يمكن اكتشاف الأخطاء انطلاقاً من المنحنى وهي المواضع التي تظهر في شكل ذروات ذات قيمة موجبة على المنحنى.



الشكل(8): مخطط للمنحنى الطاقوي لبرنامج PROSII لبنية البروتين النظري.

2.2.4.2_ قيمة النتيجة Z-Score

القيمة Z-score لبروتين ما تعرف بأنها طاقة الفصل Energy separation بين الانطواء الطبيعي ومتوسط مجموعة من النماذج النظرية المتحصل عليها. ويستعمل لتقييم النماذج المتحصل عليها بالنمذجة وذلك بمقارنة الطاقة المتحصل عليها تجريبيا مع طاقة النماذج التي تتم محاكاتها بالحاسوب للنماذج النظرية ويحسب بالعلاقة التالية (Bhattacharya et al., 2007; Li Z et al., 1998).

$$Z = \frac{E_{(\text{النموذج})} - E_{(\text{الانطواء الطبيعي})}}{\sigma_{(\text{النموذج})}}$$

حيث $E_{(\text{الانطواء الطبيعي})}$ تمثل طاقة بنية النموذج الطبيعي للبروتين، $E_{(\text{انطواء النموذج})}$ متوسط طاقة مجموعة انطواءات البنيات النظرية، و $\sigma_{(\text{نموذج})}$ تمثل الانحراف المعياري لطاقة الانطواءات النظرية (Bhattacharya et al., 2007; LiZ et al., 1998)، باستعمال $\alpha = 0.05$ و نسبة خطأ ب 5%.

3.2.4.2_الجذر التربيعي لمتوسط المربعات RMSD Score

القياسات الأولى والأكثر استعمالاً في مجال تحديد مدى التشابه البنيوي تعتمد أساساً على حساب الجذر التربيعي لمتوسط المربعات للمسافات (Root Mean Square (RMSD) Deviation بين الذرات (المحددة بالإحداثيات الكارتيزية) الموافقة لهيكل بنيتين بروتينيتين خاضعتين على توافق للمقارنة، وتعتبر قيمتها من الناحية البنيوية على مدى تشابه وتطابق بنيتين ثلاثيتي الأبعاد، حيث أنه كلما كانت قيم الـ RMSD ضعيفة، كلما زادت درجة التشابه بين البنيتين المقارنتين والعكس (Shirley et al., 2005). تؤول قيمة الـ RMSD إلى الصفر عند مقارنة بنية ما بنفسها.

وتحسب قيمة الـ RMSD حسب العلاقة التالية: إذا كان لدينا بنيتين A و B فإن قيمة الـ RMSD بينهما تحسب بالعلاقة الآتية:

$$RMSD_c = \sqrt{\frac{\sum_{i=1}^N (D(a_i, b'_i))^2}{N}}$$

حيث أن:

N- هو عدد الذرات الموضوعة الخاضعة للمطابقة بين البنيتين: الذرة b_i في البنية B موضوعة على توافق مع الذرة a_i في البنية A و القيمة $D(a_i, b_i)$ هي المسافة بين الذرات (a_i, b_i) بعد التوضّع الأمثل لجميع ذرات البنيتين التي على توافق (Mathilde et al., 2005).

النتائج والمناقشة

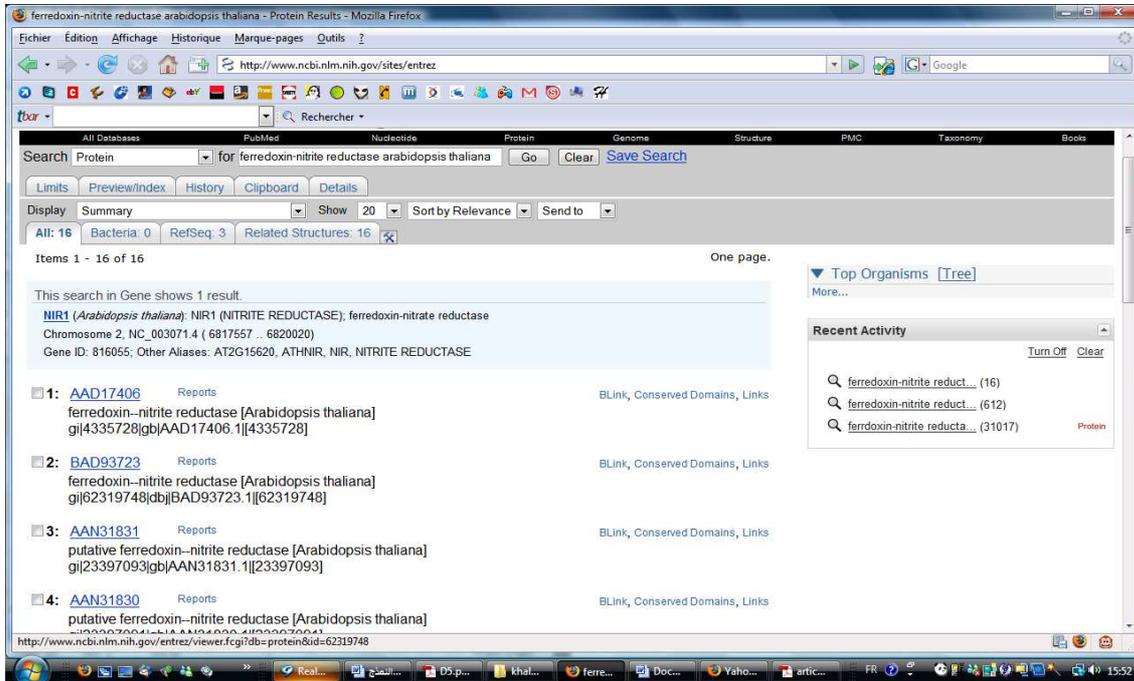
1_ النمذجة بالتمائل

إنّ عملية النمذجة بالتمائل تتألف من أربع خطوات رئيسية (Marti-Renom et al., 2000):

1.1_ القالب أو القوالب المناسبة

1.1.1_ أولا: التتابع الأميني للبروتين الهدف

تم الحصول على التتابع الكامل للأحماض الأمينية لهذا البروتين انطلاقا من موقع بنك المعطيات NCBI (National Center for Biotechnology Information) من على شبكة الانترنت www.ncbi.nlm.nih.gov/. يعد إدخال اسم الإنزيم واسم نوع النبات كموضوع بحث تحصلنا على التتابع الأميني الخاص بالـ Nitrite reductase ferredoxin- لنبات *Arabidopsis thaliana* وهو يحمل الرمز AAD17406، الذي قمنا بتنزيله بامتداد FASTA، كما هو موضح في (الشكل 1و2):



الشكل(1): نافذة العرض لموقع بنك الـ NCBI الذي يتم منه تحميل التتابع الأميني للبروتين الهدف.

PRVELKDGFFILKEKFRKGINPQEKVKIEREPMKLFMENGIEELAKKSMEELDSEKSSKDDIDVRLKWL
 GLFHRRKHQYKGFMMRLKLPNGVTTSAQTRYLASVIRKYGEDGCADVTTTRQNWQIRGVVLPDVPEILKG
 LASVGLTSLQSGMDNVRNPVGNPIAGIDPEEIVDTRPYTNLLSQFITANSQGNPDFTNLPRKWNVCVVG
 THDLYEHPHINDLAYMPANKDGRFGFNLLVGGFFSPKRCCEAIPLDWVPAADDVLP LCKAVLEAYRDLG
 TRGNRQKTRMMWLIDELGVEGFRTEVEKRMKLERGSSSEDLVKNQWERRDYFGVNPQKQEGLSFVGL
 HVPVGLRQADDMDLRLADTYGSGELRLTVEQNIIPNVETSKTEALLQEPFLKNRFSPEPSILMKGL
 VACTGSQFCGQAI IETKLRALKVTEEVERLVSVPRP IRMHWTGCPNTCGQVQVADIGFMGCLTRGEEGK
 PVEGADVYVGGRIGSDSHIGEYKKG

الشكل (2): التتابع الأميني الخاص بال Ferredoxin- nitrite reductase لنبات *Arabidopsis thaliana*

2.1.1_ثانيا: تتابعات البروتينات محددة البيئة تجريبيا الشبيهة للبروتين الهدف

بعد البحث في بنك المعطيات PDB وجدنا انه يحتوي على بنية واحدة فقط لـ Ferredoxin- Nitrite reductase والخاص بنبات السبانخ. بنيته محددة تجريبيا بواسطة الأشعة X سنة 2005 وبدقة تصميم تساوي 2.8 Å ، يحتوي على 534 حمض أميني و مجموعتين غير بروتينيتين هما مجموعة الـ Siroheme ومجموعة الـ [Fe₄-S₄] ، ووزنه الجزيئي يساوي 69291.58 (Swamy et al., 2005).

كما أن البحث الذي قمنا به أيضا سواء عن طريق برنامج Modeller او بنوك المعطيات SSEARCH-PDB (Smith et al., 1981) و قاعدة البيانات SWISSMODEL التي تستخدم عدة برامج حاسوبية كالـ BLAST و الـ HHSEARCH قادنا الى نفس النتائج المتحصل عليها سابقا حيث وجدنا أيضا أن Ferredoxin-Nitrite reductase لنبات السبانخ يأتي دائما في المرتبة الأولى كما هو موضح في (الشكل 3) من حيث نسبة التشابه والتي تبلغ 81%. وهي نسبة تشابه عالية جدا حيث تفوق بكثير النسبة 30% المعتمدة كحد أدنى لإجراء عملية النمذجة بالتماثل.

```
>P1;isoA
sequence:isoA: 0: : 0: :::-1.00:-1.00
PRVELKDGFFILKEKFRKGINPQEKVKIEREPMKLFMENGIEELAKKSMEELDSEKSSKDDIDVRLKWLGLFHRR
KHQYKGFMMRLKLPNGVTTSAQTRYLASVIRKYGEDGCADVTTTRQNWQIRGVVLPDVPEILKGLASVGLTSLQSG
MDNVRNPVGNPIAGIDPEEIVDTRPYTNLLSQFITANSQGNPDFTNLPRKWNVCVVGTHDLYEHPHINDLAYMPA
NKDGRFGFNLLVGGFFSPKRCCEAIPLDWVPAADDVLP LCKAVLEAYRDLGTRGNRQKTRMMWLIDELGVEGFRTE
EVEKRMKLERGSSSEDLVKNQWERRDYFGVNPQKQEGLSFVGLHVPVGLRQADDMDLRLADTYGSGELRLT
VEQNIIPNVETSKTEALLQEPFLKNRFSPEPSILMKGLVACTGSQFCGQAI IETKLRALKVTEEVERLVSVPRP
IRMHWTGCPNTCGQVQVADIGFMGCLTRGEEGKPEVEGADVYVGGRIGSDSHIGEYKKG.....*
```

```
>P1;2akjA
structure:2akjA: 4: : 511: :::-1.00:-1.00
RLEPRVEERDGFVWLKEEFRRSGINPAEKVKIEKDPKLFIEDGISDLATLSMEEVDKSKHNKDDIDVRLKWLGLF
HRRKHHYGRFMMRLKLPNGVTTSEQTRYLASVIKKYKGDGCADVTTTRQNWQIRGVVLPDVPEI IKGLESVGLTSL
QSGMDNVRNPVGNPLAGIDPHEIVDTRPFNTLISQFVTANSRGNLSITNLPRKWNPCVIGSHDLYEHPHINDLAY
MPATKNGKFGFNLLVGGFFSIKRCCEAIPLDWVSAEDVVPVCKAMLEAFRDLGFRGNRQKCRMMWLIDELGMEA
```

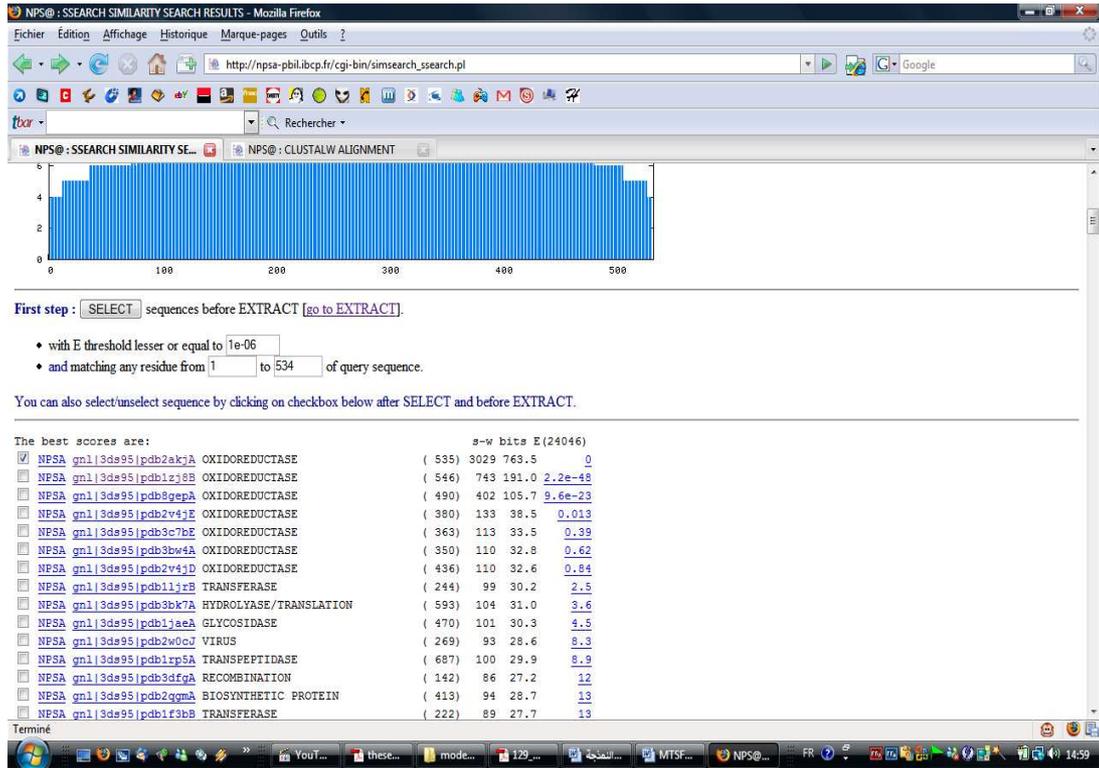
```
FRGEVEKRMPEQVLERASSEELVQKDWERREYLGVHPQKQOGLSFVGLHIPVGRLOADEMEELARIADVYGSSEL
RLTVEQNI I IPNVENSKIDSLLNEPLLKERYSPPEPILMKGLVACTGSQFCGQAI IETKARALKVTEEVQRLVSV
TRPVRMHWTGCPNSCGQVQVADIGFMGCMTRDENGKPCGADVFGGRIGSDSHLGD IYKK-----
--*
```

```
>P1;1aopA
structure:1aopA: 1: : 406: ::: -1.00: -1.00
-----
-----LLRCRLPGGVITTKQWQAIDKFAGENTIYGSIRLTNRQTFQFHG-ILP----VHQLHSVGLDALNDM
NRNVL-CTSNP---YESQLHAEAYEWAKKISEHLLP-----TYLPRKFKTTVVIPPQNDIDLHANDMNFVAI
AENGKLGFNLLVGGGLSKKTYARTASEFGYLPLEHTLVAEAVVTTQRDWNRTDRKNAKTYTLERVGVEFKA
EVERRAGTGRGDRIGWVKGIDDNWHLTLFI----ENGRILDYPARPLKTGLL-----EIAKI----HKGDFRIT
ANQNLI IPESEKAKIEKIAKESGLMNAVTPQ----RENSMACVSFPTCPLAMAEAEERFLPSFIDNIDNLMASDEH
IVMRVTGCPNGCGRAMLAEVGLV-----GKAPGRYNLHLGGRIG--TRIPRMYKE-----*
```

```
>P1;2aopA
structure:2aopA: 1: : 406: ::: -1.00: -1.00
-----
-----LLRCRLPGGVITTKQWQAIDKFAGENTIYGSIRLTNRQTFQFHG-ILP----VHQLHSVGLDALNDM
NRNVL-CTSNP---YESQLHAEAYEWAKKISEHLLP-----TYLPRKFKTTVVIPPQNDIDLHANDMNFVAI
AENGKLGFNLLVGGGLSKKTYARTASEFGYLPLEHTLVAEAVVTTQRDWNRTDRKNAKTYTLERVGVEFKA
EVERRAGTGRGDRIGWVKGIDDNWHLTLFI----ENGRILDYPARPLKTGLL-----EIAKI----HKGDFRIT
ANQNLI IPESEKAKIEKIAKESGLMNAVTPQ----RENSMACVSFPTCPLAMAEAEERFLPSFIDNIDNLMASDEH
IVMRVTGCPNGCGRAMLAEVGLV-----GKAPGRYNLHLGGRIG--TRIPRMYKE-----*
```

```
>P1;3aopA
structure:3aopA: 1: : 403: ::: -1.00: -1.00
-----
-----LRCLPGGVITTKQWQAIDKFAGENTIYGSIRLTNRQTFQFHG-ILP----VHQLHSVGLDALNDM
NRNV-----LCTSNPYSQLHAEAYEWAKKISEHLLPLPRKFKTTVVIPPQNDIDLHANDMNFVAI
AENGKLGFNLLVGGGLSKKTYARTASEFGYLPLEHTLVAEAVVTTQRDWNRTDRKNAKTYTLERVGVEFKA
EVERRAGTGRGDRIGWVKGIDDNWHLTLFI----ENGRILDYPARPLKTGLL-----EIAKI----HKGDFRIT
ANQNLI IPESEKAKIEKIAKESGLMNAVTPQ----RENSMACVSFPTCPLAMAEAEERFLPSFIDNIDNLMASDEH
IVMRVTGCPNGCGRAMLAEVGLV-----GKAPGRYNLHLGGRIG--TRIPRMYKE-----*
```

الشكل (3): التتابعات الأمينية للبنيات البروتينية المشابهة للبروتين الهدف المتحصل عليها بواسطة برنامج الـ Modeller (تحصلنا على عدد كبير من البنيات المشابهة لكن قمنا بذكر الأربعة الأولى فقط).



الشكل(4): نافذة العرض من برنامج PDB - SSEARCH الذي يعرض تتابعات البروتينات القالب حسب نسبة التشابه.

الجدول (1): يبين ترتيب البنيات البروتينية المشابهة للبروتين الهدف حسب نسبة التشابه بواسطة البرنامجين BLAST و HHSEARCH.

برنامج البحث	نسبة التشابه	دقة التصميم	رمز البروتين	مجال التشابه
BLAST	81%	2.80	2akjA	45-579
BLAST	31%	2.80	1zj8B	64-550
BLAST	31%	2.80	1zj8A	64-550
BLAST	31%	2.90	1zj9A	64-550
BLAST	31%	2.90	1zj9B	64-550
BLAST	25%	1.90	2gepA	130-585
BLAST	25%	2.40	7gepA	130-585
BLAST	25%	2.10	3aopA	131-585
BLAST	25%	1.80	4aopA	131-585
BLAST	25%	1.60	1aopA	130-585
BLAST	25%	1.75	2aopA	130-585
BLAST	25%	2.00	4gepA	130-585
BLAST	25%	2.20	5aopA	130-585

BLAST	25%	2.10	5gepA	130-585
BLAST	25%	1.80	6gepA	130-585
BLAST	25%	2.20	8gepA	130-585
BLAST	25%	2.10	3geoA	130-585
HHSEARCH	81%	2.80	2akjA	45-579
HHSEARCH	28%	2.80	1zj8A	53-576
HHSEARCH	23%	1.60	1aopA	130-575
HHSEARCH	13%	2.00	3c7bB	118-350
HHSEARCH	14%	2.00	3c7bA	314-522
HHSEARCH	24%	2.00	3c7bB	390-522
HHSEARCH	15%	2.80	1zj8A	114-346
HHSEARCH	13%	1.60	1aopA	125-342
HHSEARCH	14%	2.00	3c7bA	129-350

2.1_ مطابقة تتابع البروتين الهدف بالبروتين القالب

تحصلنا بعد إجراء عملية المطابقة ببرنامج Modeller وبرنامج CLUSTALW بين تتابع البروتين الهدف isoA و البروتين القالب 2akj على نسبة 79% و 77.99% على التوالي (الشكل 5 و 7)، حيث نلاحظ أنّ أغلب الأحماض الأمينية متطابقة بين البروتين الهدف والبروتين القالب ماعدا في بعض المناطق وهي قليلة ومتمركزة في البداية الامينية والنهاية الكاربوكسيلية.

تعتبر نسبة المطابقة المتحصل عليها عالية ومقبولة إلى حد كبير على اعتبار أنّ النسبة المقبولة كحد أدنى لإجراء عملية النمذجة بالتماثل هي 30% (Martí-Renom et al., 2000)، تقارن نسبة التتابع المتحصل عليها بدرجة التصميم (Resolution) للبنيات المتحصل عليها تجريبيا، وعليه فإنّ نسبة 79% توافقه بدرجة تصميم أقل بكثير من 3,5 Å مما يجعل بنية البروتين النموذج ذات قيمة وجودة عالية. وهذا لأنّ مصداقية طريقة النمذجة بالتماثل تعتمد بشكل مباشر على نسبة التشابه بين تتابع البروتين الهدف والقالب. بحيث إذا فاقت نسبة التشابه الـ 50% فإنّ التتابعين يعتبران متشابهين، والنمذجة ستكون عالية الجودة، والنماذج التي سنحصل عليها بعد

ذلك تشبه إلى حد كبير البنيات المتحصل عليها تجريبيا بواسطة الأشعة X ذات دقة التصميم (Resolution) أقل من 3,5 Å (Kopp et al., 2004).

أما إذا كانت نسبة التشابه بين تتابع البروتين الهدف والبروتين النموذج تتراوح من 30% إلى 50%، فإنّ النماذج التي سنحصل عليها توافق البنيات التجريبية ذات الدقة 3,5 Å (Kopp et al., 2004)، وإذا كانت نسبة التشابه أقل من 30% فإنّ النموذج سيحمل أخطاء كثيرة ومهمة (Vitkup et al 2001; Kopp et al., 2004).

```

aln.pos      10      20      30      40      50      60
2akj_lig    RLEPRVEERDGFVWLKEEFRSGINPAEKVKIEKDPMKLFIEDGISDLATLSMEEVDKSKHKNKDDIDVR
isoA        ---PRVELKDGFFILKEKFRKGINPQEKVKIEREPMKLFMENGI EELAKKSMEELDSEKSKDDIDVR
_consvrd    ****  ***  ***  *  ****  *****  *****  *  **  **  ****  *  *  *****

aln.p      70      80      90      100     110     120     130
2akj_lig    LKWLGLFHRRKHHYGRFMMRLKLPNGVTTSEQTRYLASVIKKGKDGCAVDVTRQNWQIRGVVLPDVP
isoA        LKWLGLFHRRKHQYGFMMRLKLPNGVTTSAQTRYLASVIRKYGEDGCADVTRQNWQIRGVVLPDVP
_consvrd    *****  *  *****  *****  *****  *****  *****  *****

aln.pos     140     150     160     170     180     190     200
2akj_lig    EIIKGLSVGLTSLQSGMDNVRNPVGNPLAGIDPHEIVDTRPFTNLISQFVTANSRGNLSITNLPRKW
isoA        EILKGLASVGLTSLQSGMDNVRNPVGNPIAGIDPEEIVDTRPYTNLLSQFITANSQGNPDFTNLPRKW
_consvrd    **  ***  *****  *****  *****  *****  *****  *****

aln.pos     210     220     230     240     250     260     270
2akj_lig    NPCVIGSHDLYEHPHINDLAYMPATKNGKFGFNLLVGGFFSIKRCEEAIPLDAWVSAEDVVPCKAML
isoA        NVCVVGTHDLYEHPHINDLAYMPANKDGRFGFNLLVGGFFSPKRCEEAIPLDAWVPAADDVLPCKAVL
_consvrd    *  *  *  *****  *  *  *****  *****  *****  *  *  *  *  *  *

aln.pos     280     290     300     310     320     330     340
2akj_lig    EAFRD LGFRGNRQKCRMMWLIDELGMEAFRGEVEKRMPEQVLERASSEELVQKDWERREYLGVHPQKQ
isoA        EAYRDLGTRGNRQKTRMMWLIDELGVEGFRTEVEKRMPEQVLERASSEELVQKDWERREYLGVHPQKQ
_consvrd    **  ****  *****  *****  *  *  *****  *****  *  *  *  *  *  *  *

aln.pos     350     360     370     380     390     400
2akj_lig    QGLSFVGLHVPVGRQLQADEMEELARIADVYSGELRLTVEQNIIIPNVENSKIDSLNLEPLLKERYSP
isoA        EGLSFVGLHVPVGRQLQADDMDLARLADTYSGELRLTVEQNIIIPNVETSKTEALLQEPFLKNRFS
_consvrd    *****  *****  *  ****  *  *****  *****  *****  *  *  *  *  *  *

aln.p      410     420     430     440     450     460     470
2akj_lig    EPPILMKGLVACTGSQFCGQAI IETKARALKVTEEVQRLVSVTRPVRMHWTGCPNSCGQVQVADIGFM
isoA        EPSILMKGLVACTGSQFCGQAI IETKLRLKLVTEEVERLVSVPRI RMHWTGCPNTCGQVQVADIGFM
_consvrd    **  *****  *****  *****  *****  *****  *****  *****

aln.pos     480     490     500     510     520     530
2akj_lig    GCMTRDENGKPCGADVFGGRIGSDSHLGDIIYKKA VPKDLVPVVAEILINQF GAVPR..
isoA        GCLTRGEEGKPVGADVYVGGRI GSDSHIGEIYKKG.-----.....-----
_consvrd    **  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

```

شكل (5): مقارنة تتابع البروتين النموذج مع تتابع البروتين القالب وتحديد الأحماض الأمينية المحفوظة بينهما (مشار إليها بعلامة نجمة). بواسطة برنامج Modeller.

```
>P1;2akj_lig
structureX:2akj_lig.pdb: 22 :A:+537 :A::-1.00:-1.00
RLEPRVEERDGFVWLKEEFRRSGINPAEKVKIEKDPMKLFIEDGISDLATLSMEEVDKSKHNKDDIDVRLKWLGLF
HRRKHYYGRFMMRLKLPNGVTTSEQTRYLASVIKYYKGDGCADVTTRQNWQIRGVVLPDVPFEIKGLASVGLTSL
QSGMDNVRNPVGNPLAGIDPHEIVDTRPFTNLISQFVTANSRGNLSITNLPRKWNPCVIGSHDLYEHPHINDLAY
MPATKNGKFGFNLLVGGFFSIKRCEEAIPLDAWVSAEDVVPVCKAMLEAFRDLGFRGNRQKCRMMWLIDELGMEA
FRGEVEKRMPEQVLERASSEELVQKDWERREYLVGHPQKQGLSFVGLHVPVGRLOADEMEELARIADVYSGSEL
RLTVEQNIIPNVENSKIDSLLNEPLLKERYSPPEPILMKGLVACTGSQFCGQAI IETKARALKVTEEVQRLVSV
TRPVRMHWTGCPNSCGVQVADIGFMGCMTRDENGKPCGADVFGVGRIGSDSHLGDYKKAVPCKDLVPVVAEI
LINQFGAVPR..*
```

```
>P1;isoA
sequence:isoA: : : : : 0.00: 0.00
---PRVELKDGDFILKEKFRKGINPQEKVKIEREPMKLFMENGIEELAKKSMEELDSEKSSKDDIDVRLKWLGLF
HRRKHQYQKFMRLKLPNGVTTSAQTRYLASVIRKYGEDGCADVTTRQNWQIRGVVLPDVPFEILKGLASVGLTSL
QSGMDNVRNPVGNPIAGIDPEEIVDTRPYTNLLSQFITANSQGNPDFTNLPRKWNVCVVGTHDLYEHPHINDLAY
MPANKDGRFGFNLLVGGFFSPKRCEEAIPLDAWVPAADDVLP LCKAVLEAYRDLGTRGNRQKTRMMWLIDELGVEG
FRTEVEKRMPPNGKLERGSSSEDLVKNQWERRDYFGVNPQKQEGLSFVGLHVPVGRLOADDMDLRLADTYSGSEL
RLTVEQNIIPNVETSKTEALLQEPFLKNRFSPEPSILMKGLVACTGSQFCGQAI IETKL RALKVTEEVERLVSV
PRPIRMHWTGCPNTECGVQVADIGFMGCLTRGEEGKPVGADVFGVGRIGSDSHIGEYKKG. ---.....
..-----.-...*
```

شكل (6): مطابقة بنيتي البروتينين النموذج و البروتين القالب (تمثل الفراغات المشار إليها بلون أصفر إلى مناطق عدم المطابقة (gap).

شكل (7): مطابقة تتابع البروتين النموذج مع تتابع البروتين القالب وتحديد الأحماض الأمينية المحفوظة (مشار إليها بعلامة نجمة) بواسطة برنامج CLUSTALW.

	10	20	30	40
SEQUENCE	RLEPRVEERDGFVWLKEEFRRSGINPAEKVKIEKDPMKLFIEDGISDLA			
PHD	ccccccccccccccccccccccccchhhhhhccccccccchhhhhcccccc			
isoAxx1	---PRVELKDGDFILKEKFRKGINPQEKVKIEREPMKLFMENGIEELA			
PHD	---ccccccccccccccccchhhhccccchhhhhhcccccccccccccccccc			
Prim.cons.	RLEPRVE22DGF22LKE2FR2GINP2EKVKIE22PMKLF2E2GI22LA			
	58	68	78	88
SEQUENCE	TLSMEEVDKSKHNKDDIDVRLKWLGLFHRRKHYYGRFMMRLKLPNGVT			
PHD	ccccccccccccccccchhhhhhhhhcecccccccccccccccccccccc			
isoAxx1	KKSMEELDSEKSSKDDIDVRLKWLGLFHRRKHQYKFMRLKLPNGVT			
PHD	hhhhhccccccccchhhhhhhhhcecccccccccccccccccccccccc			
Prim.cons.	22SMEE2D22K22KDDIDVRLKWLGLFHRRKH2Y2FMMRLKLPNGVT			

	106	116	126	136
SEQUENCE	TSEQTRYLASVIK	KYKGDGCADVTTRQ	NWQIRGVVLPDVPEI	IKGLES
PHD	chhhhhhhhhhhhh	ccccceeecccc	hhhccccccc	hhhhhhhhhh
isoAxx1	TSAQTRYLASVIR	KYGEDGCADVTTRQ	NWQIRGVVLPDVPEI	LKGLAS
PHD	chhhhhhhhhhhhh	ccccceeecccc	hhhccccccc	hhhhhhhhhh
Prim.cons.	TS2QTRYLASVI2	KYG2DGCADVTTRQ	NWQIRGVVLPDVPEI	2KGL2S
	154	164	174	184
SEQUENCE	VGLTSLQSGMDNVRN	PVGNPLAGIDPHEI	VDTRPF	TNLSQFVTANSR
PHD	cccccccccccccccc	cccccccccccccccc	cccccccccccc	hhhhhhhhhhhh
isoAxx1	VGLTSLQSGMDNVRN	PVGNPIAGIDPEEI	VDTRPYTNLLSQF	ITANSQ
PHD	cccccccccccccccc	cccccccccccccccc	cccccccccccc	hhhhhhhhhhhh
Prim.cons.	VGLTSLQSGMDNVRN	PVGNP2AGIDP2EI	VDTRP2TNL2SQF	2TANS2
	202	212	222	232
SEQUENCE	GNLSITNLPRKWNPC	VIGSHDLYEHPHIND	LAYMPATKNGK	FGFNLLV
PHD	cccccccccccccccc	cccccccccccccccc	cccccccccccc	cccccccccccc
isoAxx1	GNPDPITNLPRKWN	VCVGTHTDLYEHPH	INDLAYMPANKD	GRFGFNLLV
PHD	cccccccccccccccc	cccccccccccccccc	cccccccccccc	cccccccccccc
Prim.cons.	GN222TNLPRKWN2	CV2G2HDLYEHPH	INDLAYMPA2K2	G2FGFNLLV
	250	260	270	280
SEQUENCE	GGFFSIKRCEEAIP	LDAWVSAEDVVPVCK	AMLEAFRDLGFR	GNRQKCR
PHD	cccccccccccccccc	cccccccccccc	hhhhhhhhhhhh	hhhhhhhhhh
isoAxx1	GGFFSPKRCEEAIP	LDAWVPADDVLP	LCKAVLEAYRDLG	TRGNRQKTR
PHD	cccccccccccccccc	cccccccccccc	hhhhhhhhhhhh	hhhhhhhhhh
Prim.cons.	GGFFS2KRCEEAIP	LDAWV2A2DV2P2	CKA2LEA2RDLG2	RGNRQK2R
	298	308	318	328
SEQUENCE	MMWLIDELGMEAFR	GEVEKRMPEQVLER	ASSEELVQKD	WERREYLGVH
PHD	hhhhhhhhhhhhhh	hhhhhhhhhhhh	cccccccccccc	cccccccccccc
isoAxx1	MMWLIDELGVEGFR	TEVEKRMPEQVLER	GSSEDLVNKQ	WERRDYFGVN
PHD	hhhhhhhhhhhhhh	hhhhhhhhhhhh	cccccccccccc	cccccccccccc
Prim.cons.	MMWLIDELG2E2FR	2EVEKRMPE222	LER2SSE2LV2K2	WERR2Y2GV2
	346	356	366	376
SEQUENCE	PQKQGLSFVGLHIP	VGRLQADEMEELAR	IADVYGS	GELRLTVEQNI
PHD	cccccccccccccccc	cccccccccccc	hhhhhhhhhhhh	cccccccccccc
isoAxx1	PQKQGLSFVGLHIP	VGRLQADMDELAR	LADTYGS	GELRLTVEQNI
PHD	cccccccccccccccc	cccccccccccc	hhhhhhhhhhhh	cccccccccccc
Prim.cons.	PQKQ2GLSFVGLH2	PVGRLQAD2M2ELAR	2AD2YGS	GELRLTVEQNI

```

394          404          414          424
|           |           |           |
SEQUENCE    IPNVENSKIDSLNNEPLLKERYSPEPPIILMKGLVACTGSQFCGQAIIE
PHD         eccccchhhhhhhhhhhhhhhccccccccccccceeecccccccccccccc
isoAxx1    IPNVETSKTEALLQEPFLKNRFSPEPSIILMKGLVACTGSQFCGQAIIE
PHD         eccccchhhhhhhhhhhhhhhccccccccccccceeecccccccccccccc
          *****. ** .: ** : ** : ** : ** : *****. *****
Prim.cons. IPNVE2SK222LL2EP2LK2R2SPEP2IILMKGLVACTGSQFCGQAIIE

          442          452          462          472
          |           |           |           |
SEQUENCE    TKARALKVTEEVQRLVSVTRPVRMHWTGCPNSCGQVQVADIGFMGCMT
PHD         chhhhhhhhhhhhhhhhhccccceeeeeecccccccccccccccccccccc
isoAxx1    TKLRALKVTEEVEERLVSVPRIIRMHWTGCPNTCGQVQVADIGFMGCLT
PHD         chhhhhhhhhhhhhhhhhccccceeeeeeccccccccccccceeeccccccce
          ** *****. *****. ** : *****. *****. *****. *
Prim.cons.  TK2RALKVTEEV2RLVSV2RP2RMHWTGCPN2CGQVQVADIGFMGC2T

          490          500          510          520
          |           |           |           |
SEQUENCE    RDENGKPCGADVIVGGRIGSDSHLGD IYKKA VPCCKDLVPVVAEILIN
PHD         cccccccccceeeeeeccccccccceeeccccccccchhhhhhhhhhhhh
isoAxx1    RGEEGKPVGADVIVGGRIGSDSHIGEIYKKG-----
PHD         cccccccccceeeeeecccccccccccccccccc-----
          * . : ** * *****. *****. * : *****.
Prim.cons.  R2E2GKP2EGADV2VGGRIGSDSH2G2IYKK2VPCCKDLVPVVAEILIN
SEQUENCE   QFGAVPR
PHD        hcccccc
isoAxx1    -----
PHD        -----
Prim.cons.  QFGAVPR

```

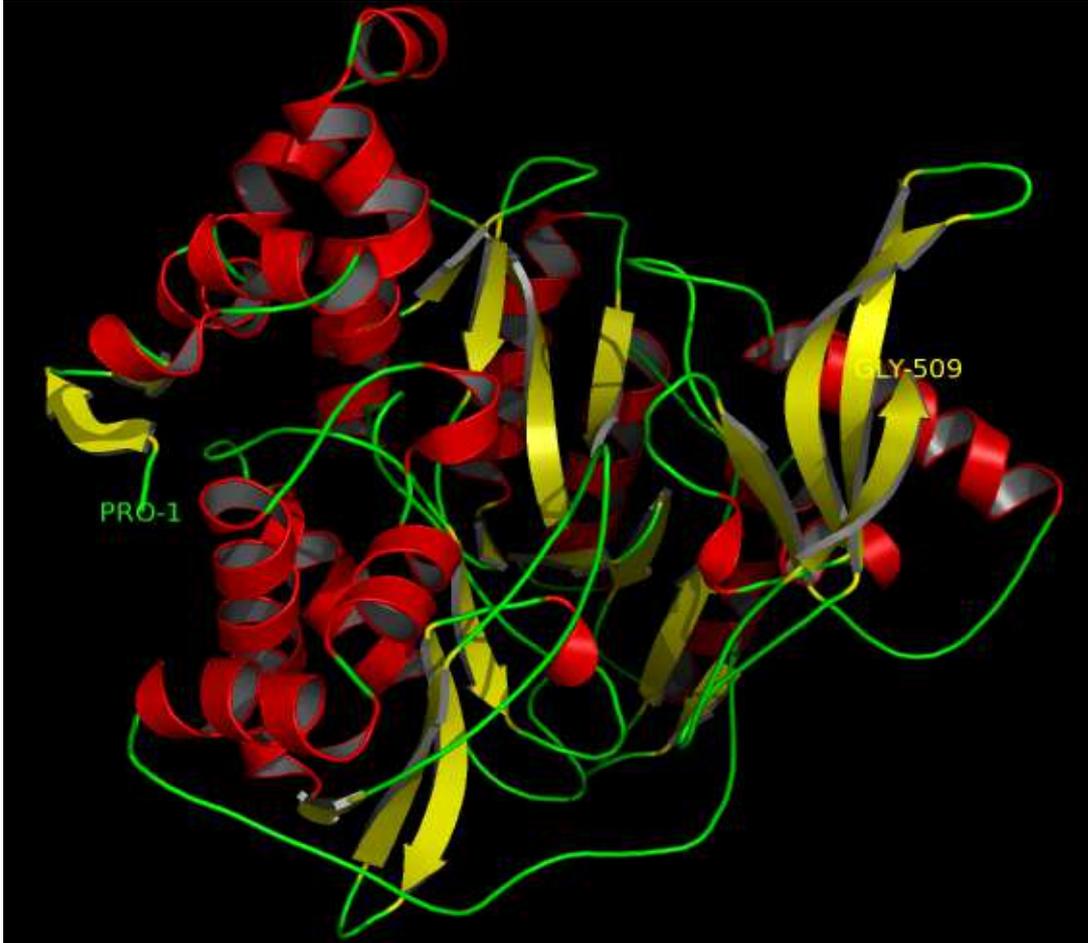
نسبة الأحماض الأمينية	عدد الأحماض الأمينية	
%77.94	417	نسبة المطابقة
%10.65	57	أطول قطعة تشابه
%3.36	18	اقصر قطعة تشابه
%8.04	43	الاختلاف التام

الجدول (2): يمثل بعض قيم المطابقة بين البروتين الهدف والقالب المتحصل عليها من برنامج

.CLUSTALW

3.1_ بناء النموذج

تم بناء 10 ملفات تعريف من نوع isoA.B999909.pdb بواسطة برنامج Modeller حيث أنّ كل ملف من بين هاته الملفات هو عبارة عن بنية ثلاثية الأبعاد للبروتين النموذج، سنختار بنية واحدة فقط لدراستها وذلك بالاستناد إلى الاستقرار الطاقوي.



الشكل (8): عرض للبنية الثلاثية الأبعاد بواسطة برنامج Pymol لنموذج ال Nitrite reductase المتحصل عليه، حيث يبين البنيات الثانوية باللون الأحمر (حلزون α) واللون الأصفر (وريات β) ومناطق الانعطاف باللون الأخضر.

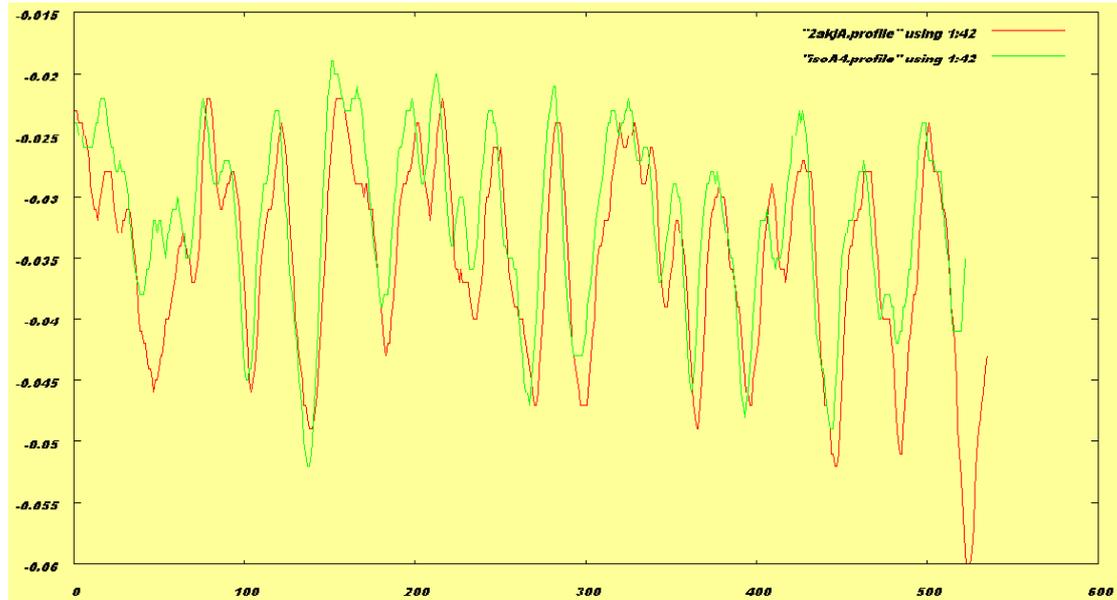
4.1_ تقييم النموذج النظري المتحصل عليه

النموذج المتحصل عليه عن طريق النمذجة بالتماثل يجب أن يتم تقييمه من عدة نواحي: الكيمياء الفراغية Stereochemistry، طاقة البروتين Protein energy ، وكذا حساب قيمة التوافق RMSD بين الهيكل الكربوني للنموذج والقالب.

1.4.1_ حساب قيمة DOPE

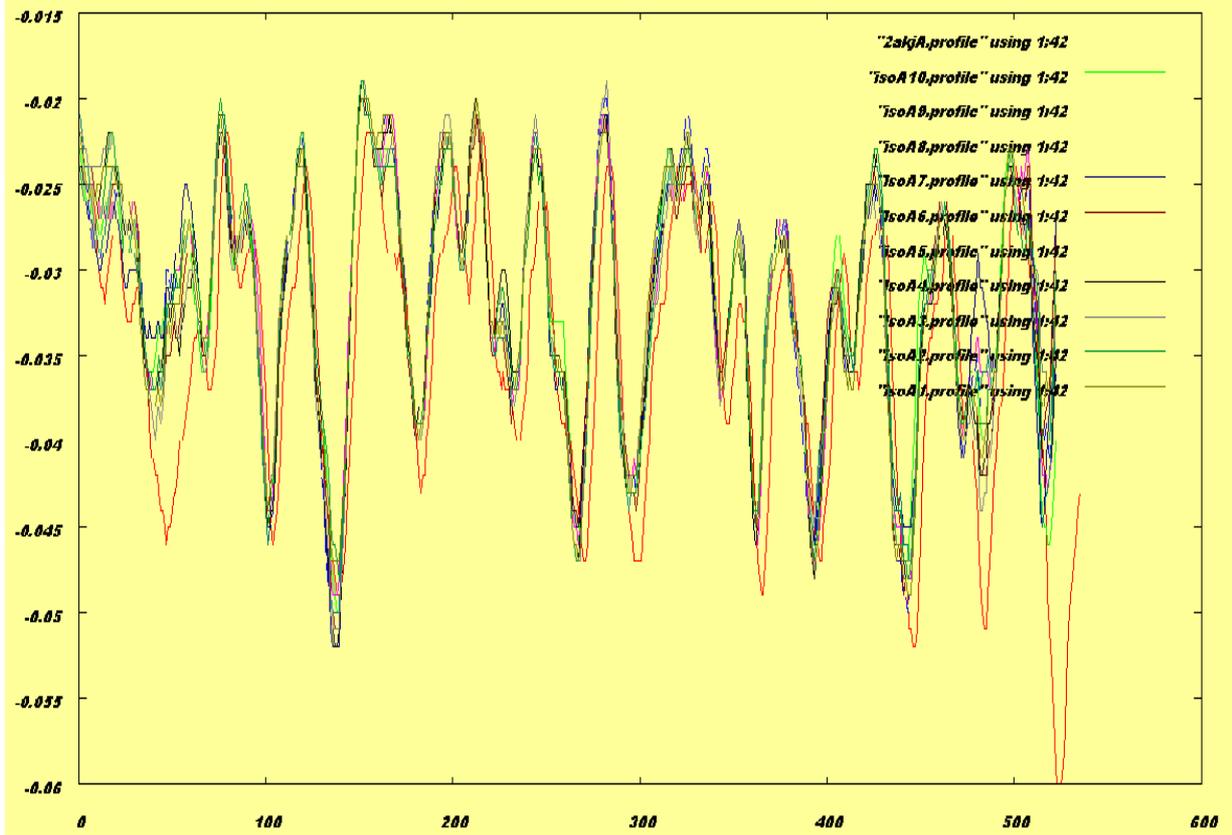
يعتمد في معرفة استقرار وثبات الجزيئات على تقييم طاقتها الداخلية، حيث أن الجزيئات العشرة المتحصل عليها للنموذج المتنبأ بها متقاربة الطاقة، وهي قريبة أيضا من قيمة ال DOPE لبنية القالب، وذلك بعد قياس (DOPE) Discrete Optimized ProteinEnergy بواسطة برنامج ال Modeller كما في (الجدول 3). بعد ذلك قمنا باختيار الجزيء الرابع من النماذج المتحصل عليها للدراسة، باعتباره أحسن هذه النماذج من حيث الاستقرار الطاقوي وكذا استنادا إلى التوافق الكبير بين المنحنى الطاقوي له والمنحنى الطاقوي للبروتين القالب 2akj كما هو موضح في (الشكل 9).

وهذا ما يدل على أنّ طاقة النموذج المتحصل عليه قريبة جدا من طاقة البروتين القالب المحدد تجريبيا أي أنّ استقرارها الطاقوي متقارب، وهذا ما يرفع من قيمة النمذجة بالتماثل ويحاكيها بالتجارب المخبرية.



شكل (9) : منحنين بيانيين يمثلان مقارنة بين قيمتي DOPE للنموذج المتحصل عليه (isoA.B9999004.pdb) باللون الأخضر و القالب 2akj باللون الأحمر.

ويمثل (الجدول 4) بعض الخصائص البنوية و الطاقوية لهذا الجزيء، كما تمثل المنحنيات الممثلة في (شكل 10) مقارنة قيمة DOPE للبروتين قالب (لون أحمر) مع النماذج المتحصل عليها (IsoA1, ..., IsoA10)، حيث نلاحظ تطابق كبير بين النماذج المتحصل عليها بالنمذجة بالمماثلة والقالب، وهذا مايعطي لنا فكرة واضحة على الاستقرار الطاقوي لجزيئات البروتينات المتحصل عليها، باعتبارها موافقة لمقدار الطاقة للجزيئات المتحصل عليها مخبريا.



شكل (10): منحنيات بيانية تمثل مقارنة بين قيم DOPE للنماذج العشرة المتحصل عليها (isoA10.pdb.....isoA1.pdb) الممثلة بألوان مختلفة و القالب 2akj بالون الأحمر.

جدول(03): يمثل القيم المثلى لطاقة كل جزيء بروتين من النماذج المتحصل عليها بواسطة برنامج

.Modeller

اسم ملف بنية النموذج المتحصل عليه	قيمة DOPE score
isoA.B99990001.pdb	-57401.35156
isoA.B99990002.pdb	-57333.39453
isoA.B99990003.pdb	-57736.35547
isoA.B99990004.pdb	-57984.07422
isoA.B99990005.pdb	-57709.65234
isoA.B99990006.pdb	-57668.04688
isoA.B99990007.pdb	-57753.37109
isoA.B99990008.pdb	-57571.63672
isoA.B99990009.pdb	-57626.10938
isoA.B99990010.pdb	-57568.52344

جدول (04): بعض الخصائص البنوية و الطاقوية للنموذج الرابع من Nitrite reductase متحصل عليها

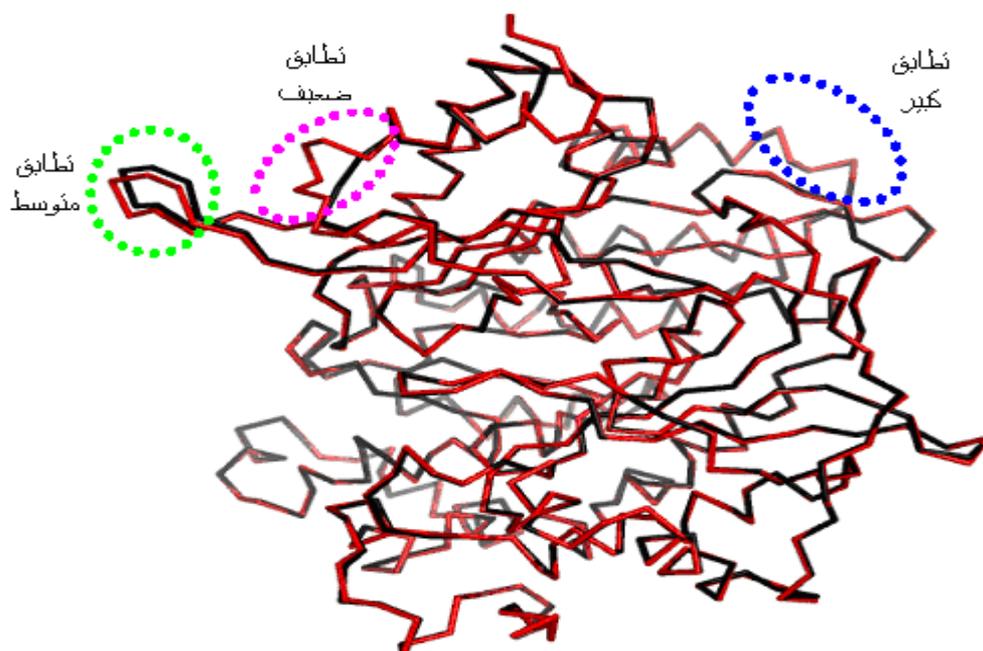
بواسطة برنامج Modeller.

sequence identity	79
Sequence length	509
Compactness	0.327395
Native energy (pair)	-420.342702
Native energy (surface)	-26.093330
Native energy (combined)	-9.954364
Z score (pair)	-11.663129
Z score (surface)	-9.028343
Z score (combined)	-13.346562

2.4.1_ نتيجة المطابقة باستعمال الجذر التربيعي لمتوسط المربعات RMSD Score

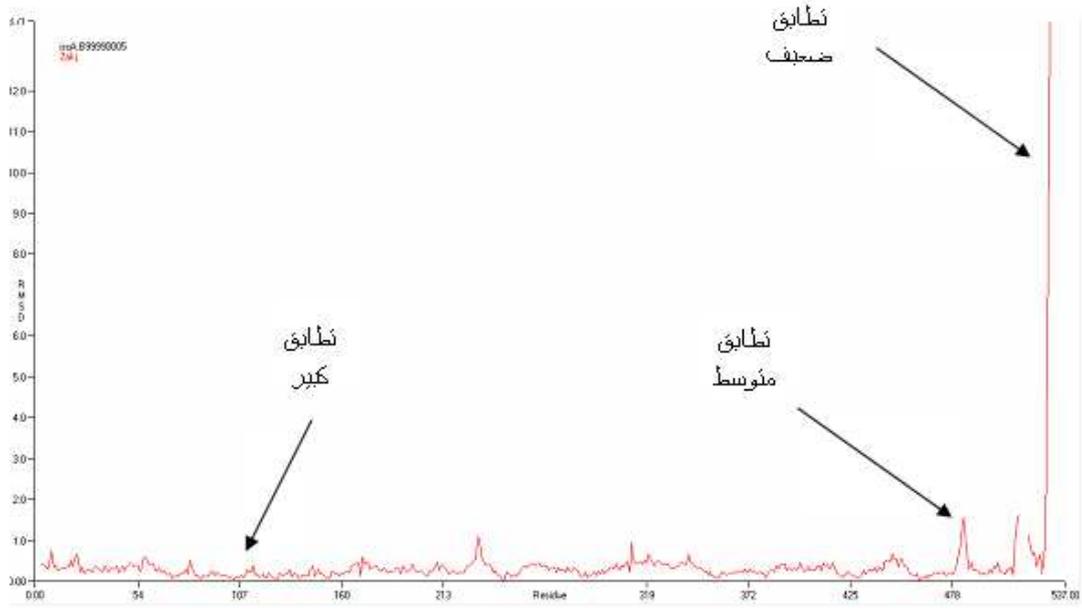
تبين قيمة الـ RMSD بين بروتينين اثنين درجة التطابق بين هيكل السلسلتين الكربونيتين α C لهما (Bino et al., 2003). باستعمال برنامج الـ Pymol تحصلنا على قيمة صغيرة جدا للـ RMSD و التي قدرت بحوالي $0,2\text{\AA}$ ، كما هو موضح في (الشكل 11). حيث تدل القيم المنخفضة أو الصغيرة لـ RMSD على المطابقة الجيدة وعلى تشابه البنيتين والعكس بالعكس (Mathilde et al., 2005; Turk et al., 1991; Ye et al., 2003).

وهذا ما يوضح أنّ هيكلي السلسلتين الرئيسيتين للبروتين النموذج والقالب متطابقتان بشكل كبير في أغلب المناطق، بينما لاحظنا أن هناك تطابق ضعيف ومتوسط في مناطق قليلة جدا، أغلبها تقع في مناطق الانعطاف والنهايات المفتوحة (الأمينية والكربوكسيلية)، وهي بعيدة عن الموقع الفعال مما يقلل من تأثيرها، وهذا ما يزيد من مصداقية بنية النموذج المتحصل عليها ويقربها ببنية القالب المحددة تجريبيا.



شكل (11): يمثل مدى المطابقة بين بنية البروتين القالب باللون الأسود و بنية النموذج المتحصل عليه باللون الأحمر بواسطة البرنامج Pymol.

يمثل المنحنى (الشكل 12) قيمة الـ RMSD بين بنية النموذج المتحصل عليه والقالب 2akj التي تم حسابها بواسطة برنامج الـ VMD، حيث نلاحظ أنّ أغلب قيم الـ RMSD محصورة بين القيمة 0Å و القيمة 1Å في معظم مناطق البروتين ماعدا بعض المناطق فقط كالنهاية الكربوكسيلية أين لا يوجد أي تطابق بينهما وهي توافق النتائج المتحصل عليها بواسطة برنامج Pymol.



الشكل (12): نتائج قياس RMSD بين بنية البروتين القالب والنموذج بواسطة البرنامج الحاسوبي VMD.

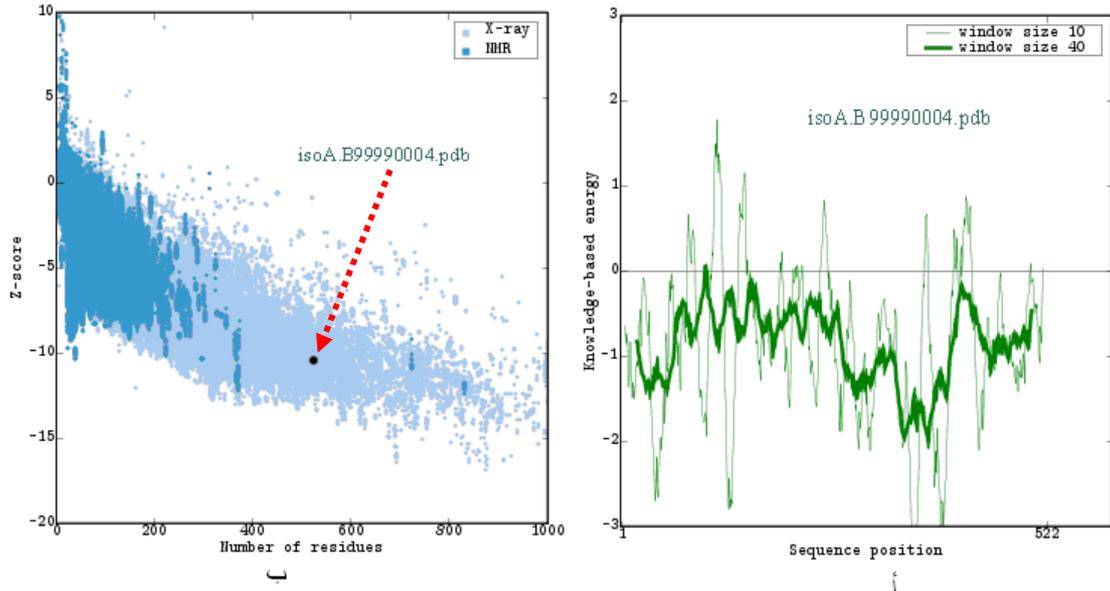
3.4.1_ تقييم البروتين النموذج المتحصل عليه بواسطة برنامج PROSAII

يقارن برنامج PROSAII قيم Z scores لبنية نموذج البروتين الهدف المتحصل عليها بقيم Z scores لبنية البروتين القالب. حيث أن قيمة Z scores للبنية البروتينية بشكل عام هو قياس مدى الملائمة بين تتابع أحماضها الأمينية وبنيتها الثلاثية.

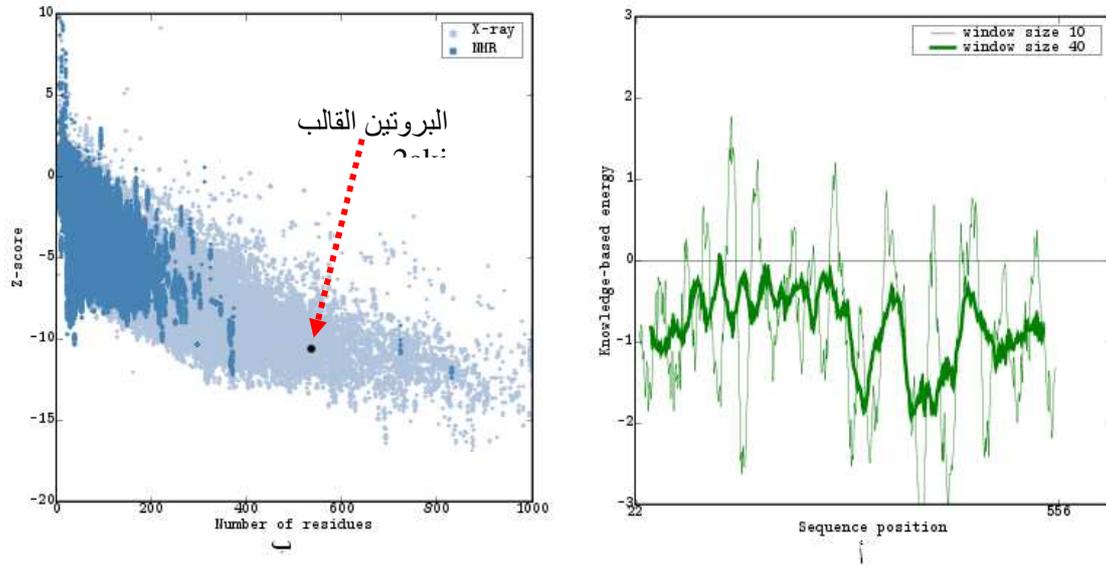
يمثل (الشكل 13) مواضع لقيم Z scores للبروتينات محددة البنية في بنك PDB تجريبيا بواسطة الرنين النووي المغناطيسي NMR على شكل نقاط باللون الأزرق الداكن وكذا البروتينات المحددة البنية بواسطة الأشعة X باللون الأزرق الفاتح، حيث نلاحظ أن قيمة Z scores للنموذج المتحصل عليه (isoA.B99990004.pdb) والمساوية لـ -10.37 تقع ضمن مجال البروتينات المحددة بواسطة الأشعة X، كما أن هذه القيمة قريبة جدا من قيمة Z scores لبنية البروتين القالب المحدد تجريبيا (-10.54) وهذا ما يدل على أن التوضع الفراغي للأحماض الامينية للبنية المتحصل عليها ذات دقة عالية تقارب 2.8.

كما أن المنحنى الطاقوي الممثل باللون الأخضر الداكن (الشكل 13-أ) لبنية النموذج المتحصل (isoA.B99990004.pdb) عليه بواسطة برنامج PROSAII يؤكد أن أغلب قيمه تقع في المجال السالب ماعدا في منطقة واحدة (لها قيمة موجبة)، كما نلاحظ أنه قريب جدا من المنحنى الطاقوي لبنية القالب وهذا أيضا يؤكد الحالة الطاقوية الجيدة لبنية النموذج المتحصل عليه ويزيد

من مصداقيته، الطاقة السالبة تدل على الاستقرار الطاقوي الناتج من التوضع الصحيح للروابط الناتجة من ارتباط ذرات الاحماض الامينية للبروتين



شكل(13): (أ) المنحنى الطاقوي للنموذج النظري المتحصل عليه بالنمذجة .(ب) مخطط Z-Score (تمثل البقعة بلون ازرق داكن وأزرق فاتح قيم Z-Score لكل من البروتينات محددة البنية بالرنين المغناطيسي والأشعة السينية على التوالي).



شكل(14): (أ) المنحنى الطاقوي للبروتين القالب 2akj.(ب) مخطط Z-Score (تمثل البقعة بلون ازرق داكن وأزرق فاتح قيم Z-Score لكل من البروتينات محددة البنية بالرنين المغناطيسي والأشعة السينية على التوالي).

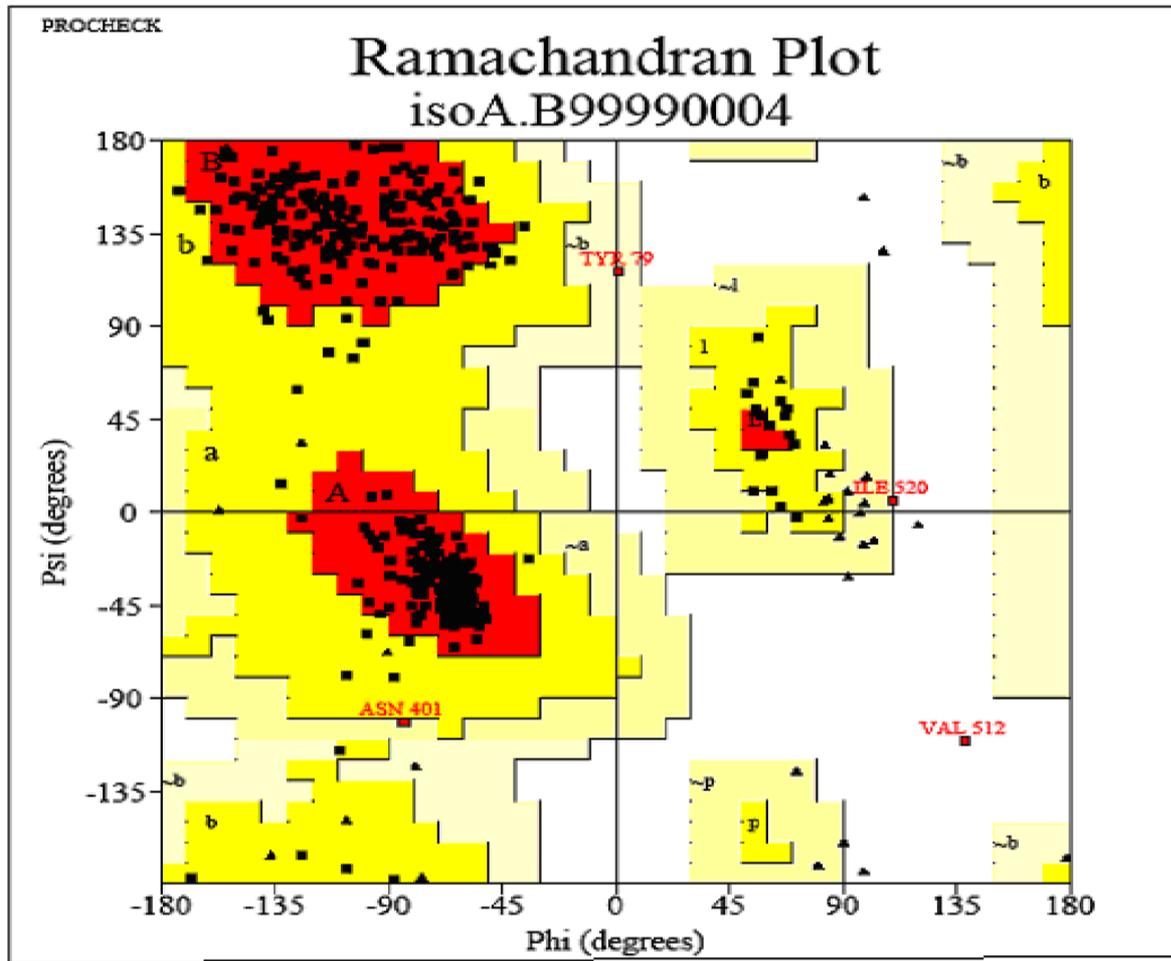
4.4.1_ نتائج مخطط Ramachandran plot للبروتين النموذج المتحصل عليه

تم تمثيل تغيرات قيم زوايا الدوران ϕ و ψ للشكل الفراغي للسلاسل الجانبية الأكثر استقرارا في رسم بياني ثنائي البعد عرف باسم مخطط Ramachandran Plot (Ramachandran et al., 1963). ليست كل قيم ϕ و ψ ممكنة دون حدوث تصادم بين الذرات، بحيث أنّ المنطقة الأكثر ملائمة هي باللون الأحمر، بينما اللون الأصفر الداكن يمثل المناطق المسموح بها، واللون البني الفاتح يمثل المناطق المسموح بها بدرجة أقل، بينما اللون الأبيض يمثل المناطق غير المسموح بها.

تكون قيم زوايا الدوران التجريبية المتحصل عليها بطريقة حيود الأشعة السينية X أو الرنين النووي المغناطيسي NMR دوما في المناطق المسموحة من المخطط ما عدا الغلايسين (الذي لا يملك سلسلة جانبية) ، أو البرولين (حمض أميني يستطيع تشكيل رابطة ببتيدية من نوع Cis).

بالنسبة للنموذج الذي قمنا ببنائه فإن النتائج التحليلية لتوزع أحماضه الأمينية على المناطق السابقة ممثلة في مخطط الـ Ramachandran plot، المتحصل عليه بواسطة البرنامج PSVS (Bhattacharya et al., 2007) كما هو موضح في (الشكل 15) و (الجدول 5).

حيث كانت نسب الأحماض الأمينية الممثلة على مخطط الـ Ramachandran plot كالأتي: المتوضعة في المنطقة الأكثر ملائمة هي 89,9% وهي نسبة عالية جدا، وهي تدل على أنّ أغلب قيم ϕ و ψ للأحماض الأمينية للبنية المتحصل عليها سليمة ومتوضعة بشكل جيد، بينما قيم ϕ و ψ للأحماض الأمينية التي تنتمي إلى المناطق المسموحة نسبتها 9,2%، والتي تخص الأحماض التي تنتمي إلى المناطق المسموح بها بدرجة أقل والمناطق غير المسموح بها نسبتها 0,5% لكل منهما، أي حوالي حمضين أميين فقط هي التي تنتمي إلى المناطق غير المسموح بها، وهذا ما يعطي للبنية المتحصل عليها مصداقية عالية تقارب ما يتحصل عليه عند تحديد البنية بالأشعة X أو بالرنين النووي المغناطيسي NMR (Bhattacharya et al., 2007). وتؤكد النتائج الطاقوية السابقة والمتعلقة بالتوضع الفراغي لذرات البروتين.



شكل (15): مخطط Ramachandran يبين توزيع الأحماض الأمينية حسب قيم الزاويتين ϕ و ψ .

يمكن تلخيص كل قيم توزيع الأحماض الأمينية في (الجدول 5):

جدول (05): نسبة توزيع الأحماض الأمينية وفق مخطط Ramachandran plot.

الأحماض الامينية		مناطق المخطط
النسبة	العدد	
%89,9	399	المنطقة الأكثر ملائمة
%9,2	41	المناطق المسموحة
%0,5	2	المناطق المسموح بها بدرجة اقل
%0,5	2	المناطق غير المسموح بها
%100	444	الأحماض الأمينية عدا الغلايسين والبرولين

2_ الدراسة التحليلية لبنية Nitrite reductase

1.2_ البنية الأولية لـ Nitrite reductase

البنية الأولية للبروتين هي عبارة عن تسلسل لتتابع الأحماض الأمينية في سلسلة خطية تدعى بمتعدد الببتيد، والتي يمكن صياغتها كتتابع لأحرف أبجدية $\sum aa$ ، بحيث يرمز لكل حمض أميني بحرف واحد أو ثلاثة حروف الأولى من اسمه، و يمكن الحصول وتحميل البنية ثلاثية الأبعاد لكل حمض أميني من بنوك المعطيات على الانترنت (Boeckmann et al., 2003). بالنسبة للبنية الأولية لـ Nitrite reductase هي عبارة عن سلسلة ببتيدية واحدة، تتكون من 509 حمض أميني و بوزن جزيئي يقدر ب 57220.5 غ/مول مرتبطة خطيا بواسطة روابط ببتيدية وقد تمكنا بواسطة البرنامج الحاسوبي ProtParam (Gasteiger et al., 2005) من التحليل الدقيق لهذه السلسلة الببتيدية حيث كانت النتائج كالاتي:

10	20	30	40	50	60
PRVELKDGFF	ILKEKFRKGI	NPQEKVKIER	EPMKLFMENG	IEELAKKSME	ELDSEKSSKD
70	80	90	100	110	120
DIDVRLKWLG	LFHRRKHQYG	KFMMRLKLPN	GVT TSAQTRY	LASVIRKYGE	DGCADVTTRQ
130	140	150	160	170	180
NWQIRGVVLP	DVPEILKGLA	SVGLTSLQSG	MDNVRNPVGN	PIAGIDPEEI	VDTRPYTNLL
190	200	210	220	230	240
SQFITANSQG	NPDFTNLPRK	WNVCVVGTHD	LYEHPHINDL	AYMPANKDGR	FGFNLLVGGF
250	260	270	280	290	300
FSPKRCEEAI	PLDAWVPADD	VLPLCKAVLE	AYRDLGTRGN	RQKTRMMWLI	DELGVEGFRT
310	320	330	340	350	360
EVEKRMPNGK	LERGSSEDLV	NKQWERRDYF	GVNPQKQEGL	SFVGLHVPVG	RLQADDMDEL
370	380	390	400	410	420
ARLADTYGSG	ELRLTVEQNI	IIPNVETSKT	EALLQEPFLK	NRF SPEPSIL	MKGLVACTGS
430	440	450	460	470	480
QFCGQAIET	KLRALKVTEE	VERLVSVPRP	IRMHWGTGCPN	TCGQVQVADI	GFMGCLTRGE
490	500	509			0
EGKPVEGADV	YVGGRIGSDS	HIGEIYKKG			

الشكل (15): التابع الأميني المفصل لـ Nitrite reductase بواسطة برنامج ProtParam

الجدول (06): يمثل كل التفاصيل التركيبية الدقيقة للـ Nitrite reductase بواسطة برنامج .ProtParam

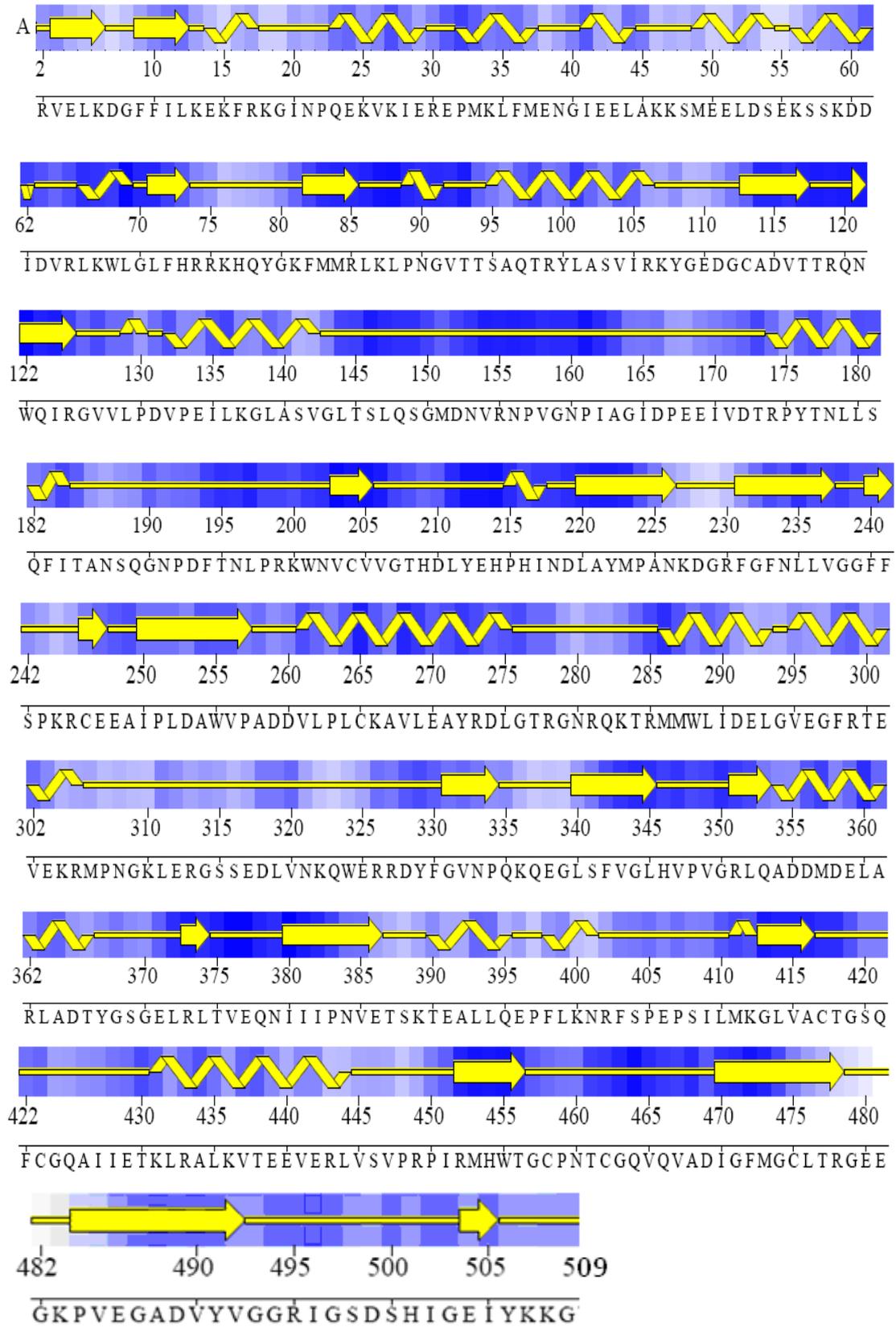
509	العدد الإجمالي للأحماض الأمينية
57220.5	الوزن الجزيئي الكلي للبروتين
71	عدد الأحماض الأمينية موجبة الشحنة (Asp + Glu)
69	عدد الأحماض الأمينية سالبة الشحنة (Arg + Lys)
C ₂₅₃₁ H ₄₀₃₃ N ₇₁₃ O ₇₅₂ S ₂₃	الصيغة العامة للبروتين
8052	العدد الإجمالي للذرات المكونة للبروتين
2531	العدد الإجمالي لذرات الكربون C
4033	العدد الإجمالي لذرات الهيدروجين H
713	العدد الإجمالي لذرات النيتروجين N
752	العدد الإجمالي لذرات الأكسجين O
23	العدد الإجمالي لذرات الكبريت S

2.2_ البنية الثانوية للـ Nitrite reductase

البنية الثانوية للبروتين ناتجة عن انطواء ذاتي للسلسلة الببتيدية بسبب تفاعلات الكترولستاتيكية، وهي مستقرة بواسطة روابط هيدروجينية منتظمة، والتي تتشكل بين المجموعات القطبية C=O و N-H حيث يمكن حسب وضعيتها تمييز ثلاث أنماط من البنيات المستقرة بواسطة الروابط هيدروجينية. أكتشفت البنية الثانوية عام 1951م بواسطة Pauling و معاونوه قبل ظهور الطرق التجريبية وذلك بالاعتماد على المعايير الفيزيائية والحيومترية (Pauling et al., 1951 ; Eisenberg et al., 2003)، حيث حددا عدة خصائص بنيوية لعدة أنماط بروتينية وهي (الحلزون α ، الوريقات β ، مناطق الانعطاف).

ولتوضيح البنيات الثانوية المكونة لإنزيم الـ Nitrite reductase بصورة بسيطة وواضحة، قمنا بدراسة نموذج البنية ثلاثية الأبعاد للبروتين النموذج المتحصل عليها isoA000004 بواسطة البرنامج الحاسوبي PROCHECK-NMR (Laskowski et al., 1996)، حيث تحصلنا على نموذج بسيط ثنائي البعد يوضح البنيات الثانوية المكونة للبروتين Model secondary structures بشكل بسيطها ويسهل من عدها وتنظيم ترتيبها، كما هو موضح في (الشكل 16).

Model 1.



الشكل(16):يبين صورة توضيحية للبنيات الثانوية المكونة لل Nitrite reductase بواسطة برنامج

.PROCHECK-NMR

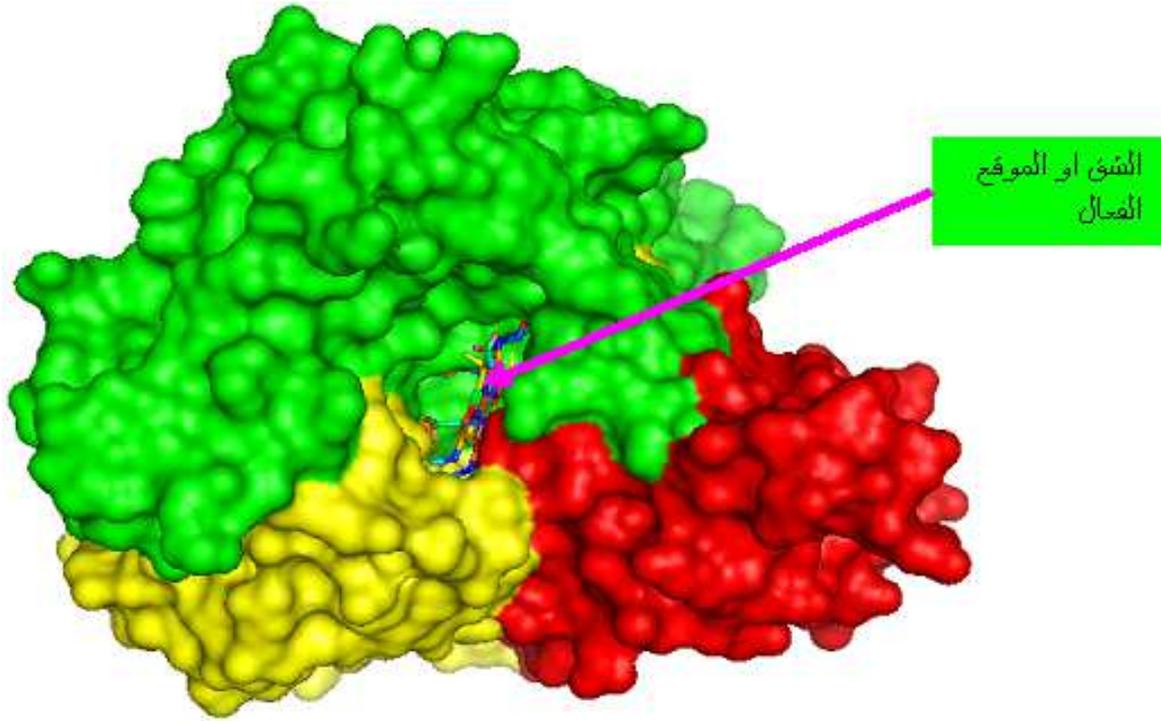
3.2_ دراسة البنية الثالثة للـ Nitrite reductase بواسطة برنامج العرض Pymol

تتكون البنية الثالثة للبروتين من تجمع عدد من البنيات الثانوية الذي ينتج عنه شكل ثلاثي الأبعاد، يضمن ثبات هذه البنية وجود مجموعة من الروابط المتشكلة بين عناصر الأحماض الأمينية فزيادة عن الروابط الهيدروجينية نجد (الجسور ثنائية الكبريت، الروابط الملحية و تجاذب السلاسل الكارهة للماء). بعد عدة انطواءات التي تخضع لها السلسلة الببتيدية، ينطوي البروتين ككل تلقائيا ليأخذ في الأخير شكلا نهائيا يحدد تخصصا وظيفيا حيويا معيناً، يحدد هذا الانطواء أساسا نتيجة طمر الأحماض الأمينية الكارهة للماء نحو الداخل مما يعطي البروتين شكلا كرويا مضغوطا، كما أنّ ثبات البنية الثالثة له أيضا علاقة بالشروط الخارجية كدرجة الحرارة ودرجة الحموضة.

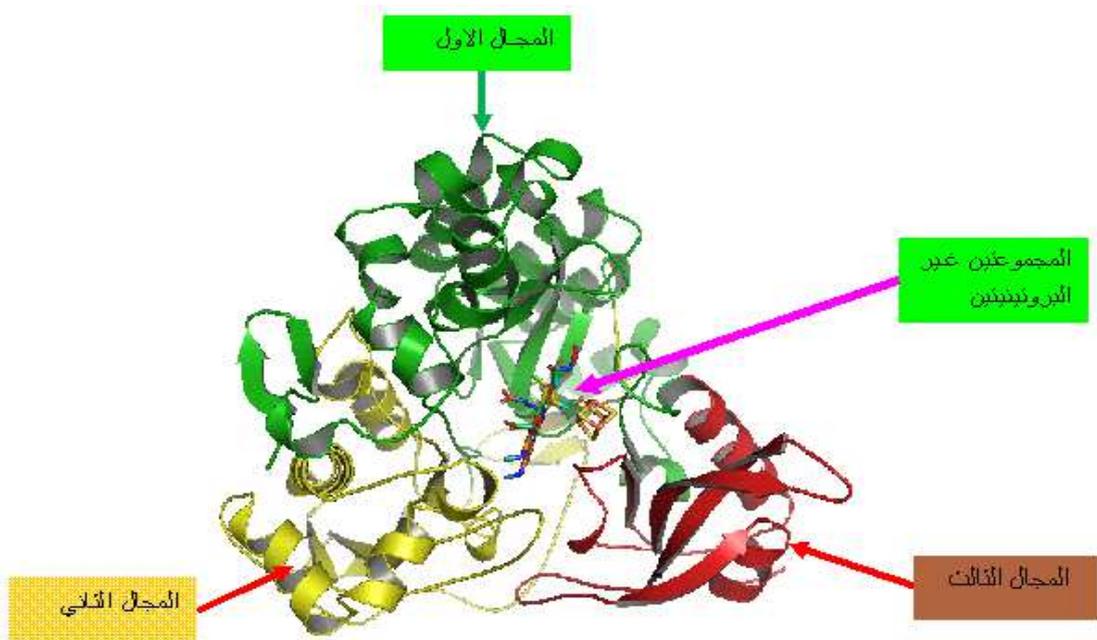
يمكن أن نميز في تركيب البنية الثالثة تجمع خاص للبنيات الثانوية لتشكيل ما يسمى بالنمط Motifs، مثل البنية البسيطة التي يطلق عليها دبوس الشعر (hairpin) β و α . أو بصورة أكثر تعقيدا كالنمط المسمى المفتاح الإغريقي (Greek key) β - α - β ، تتواجد هذه الأنماط Motifs في عدد كبير من البروتينات أين يكون لها دور بنيوي وحتى وظيفي. عدة أنماط مختلفة يمكنها التجمع مع بعض في وضع كروي مضغوط يدعى المجال Domain. هذا الأخير يمكن أن يكون مستقل بنيويا وحتى وظيفيا.

بالنسبة لإنزيم Nitrite reductase لنبات *Arabidopsis thaliana* فإنه يتكون من سلسلة واحدة متعددة الببتيد التي تتألف بدورها من مجموعة من البنيات الثانوية (حلزون α ، الوريقات β ، مناطق الانعطاف) والتي تنطوي لتشكيل بنية كروية تدعى بالبنية الثالثة، حيث يمكن تقسيمها استنادا إلى بنية البروتين المستعمل كقالب إلى ثلاث مجالات 3 Domain (المجال الأول، المجال الثاني و المجال الثالث).

أما البنية العامة فهي ذات انطواء كروي تحتوي على جيب أو شق عميق به مجموعتين غير بروتينيتين ال $[Fe_4S_4]$ و Siroheme وهو الموقع الفعال للـ Nitrite reductase كما هو موضح في (الشكل 17،18) .



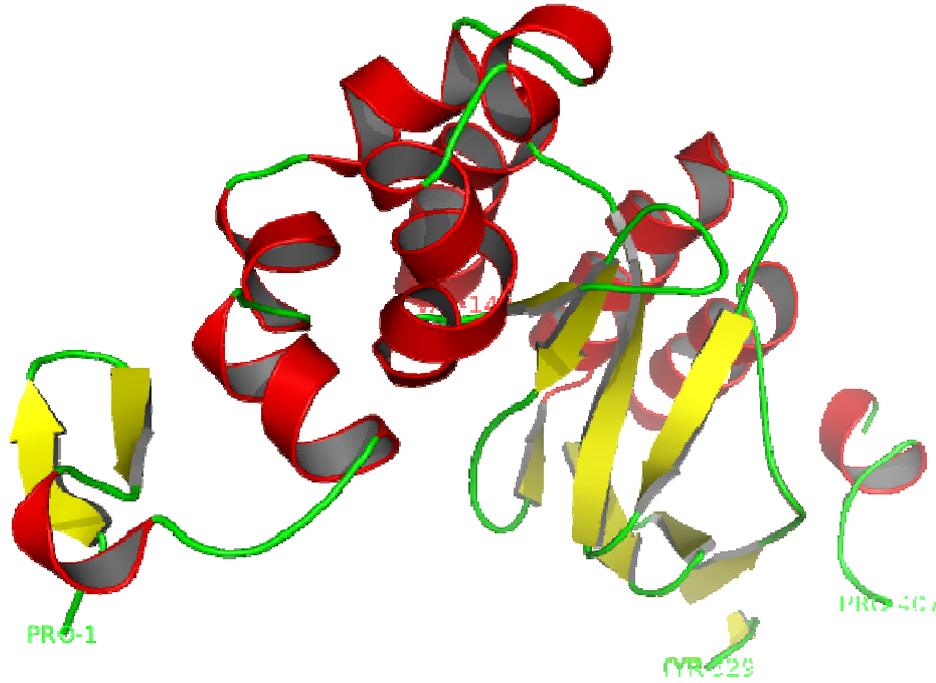
الشكل (17) : عرض البنية الثلاثية الأبعاد بالشكل المجسم العام لبنية نموذج Nitrite reductase المتحصل عليه، بواسطة برنامج Pymol.



الشكل (18) : عرض البنية الثلاثية الأبعاد بالشكل الشريطي لبنية نموذج Nitrite reductase المتحصل عليه وهو يوضح المجالات الثلاث له، بواسطة برنامج Pymol.

1.3.2_ المجال الأول Domain1

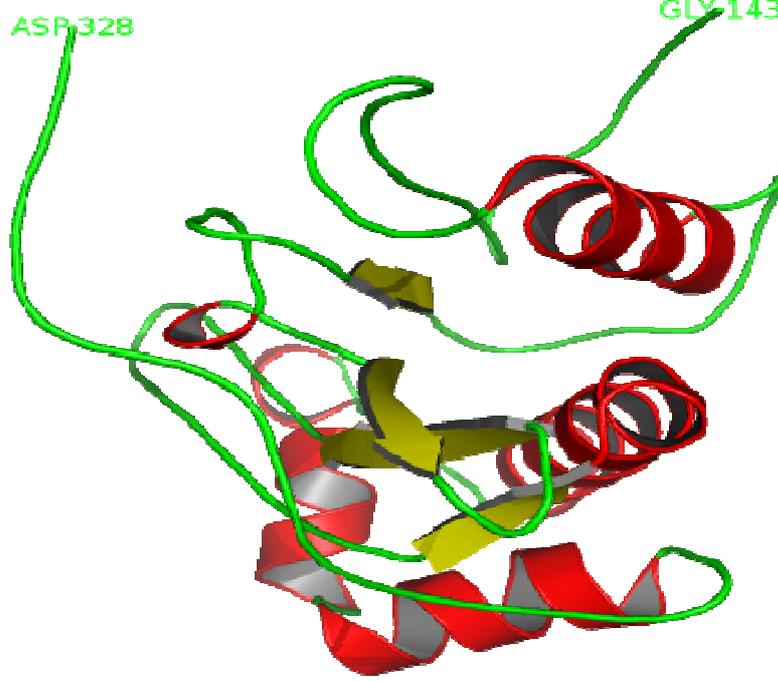
يتألف المجال الأول من 220 حمض أميني (من الحمض الأميني 1 إلى 142 ومن الحمض الأميني 329 حتى الحمض الأميني 407)، حيث يبدأ بقطعة صغيرة (من الحمض الأميني الأول 1 إلى الحمض الأميني 69) والتي تتكون من وريقتين β sheet متعاكستين، تليها خمس قطع حلزونية صغيرة α -helice. ثم تتبع بقطعة أكبر تتكون من 4 قطع α -helice و 9 قطع β sheet، القطعة الأخيرة يمكن تقسيمها إلى تحت مجالين متناظرين. تحت المجال الأول يبدأ من الحمض الأميني رقم 70 حتى الحمض الأميني رقم 142 وهو يتكون من قطعتين α -helice و 4 قطع β sheet. أما تحت المجال الثاني فيتكون أيضا من 4 قطع β -sheet وقطعتين α -helice حيث يبدأ من الحمض الأميني رقم 329 حتى الحمض الأميني رقم 407.



الشكل (19) : عرض البنية الثلاثية الأبعاد بالشكل الشريطي للمجال الأول لـ Nitrite reductase بواسطة برنامج Pymol.

2.3.2_ المجال الثاني

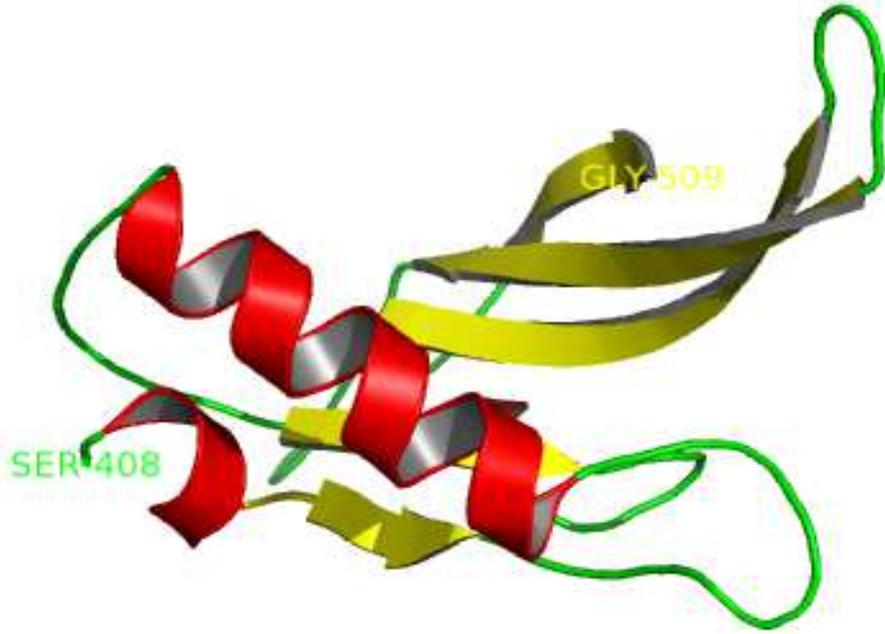
يمتد من الحمض الأميني رقم 143 حتى الحمض الأميني رقم 328، حيث يحتوي على أربعة أوراق مطوية مركزية β -sheet متعاكسة، محاطة بأربعة قطع من نوع α -helice كما له عروات طويلة نوعا ما مقارنة بالمجالين الآخرين.



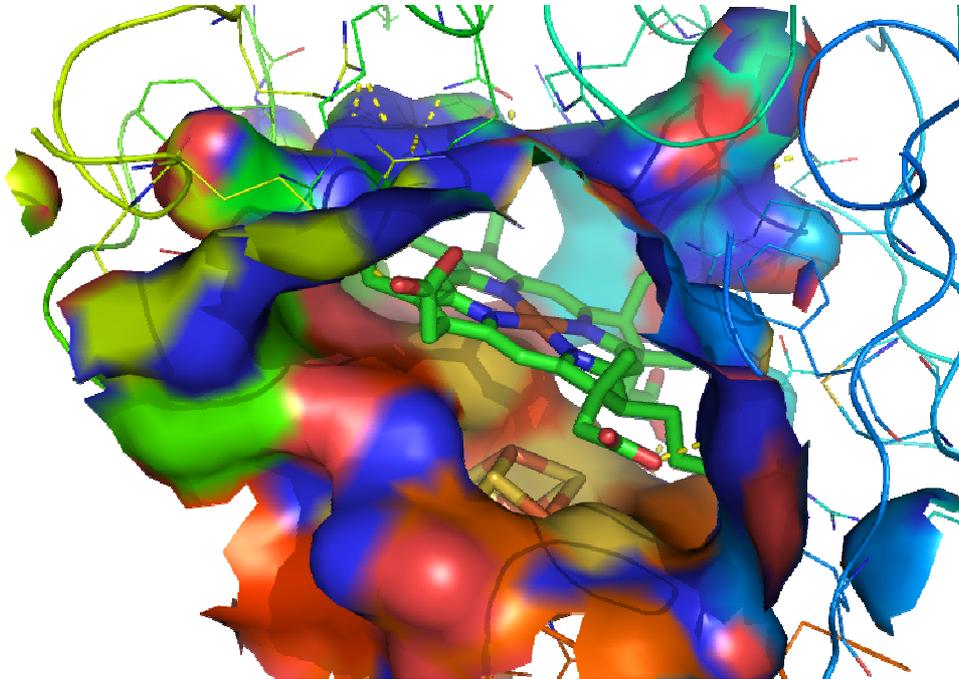
الشكل (20) : عرض البنية الثلاثية الأبعاد بالشكل الشريطي للمجال الثاني لـ Nitrite reductase بواسطة برنامج Pymol.

3.3.2_ المجال الثالث

يمتد من الحمض الأميني رقم 408 حتى الحمض الأميني الأخير، حيث يحتوي على خمس قطع من نوع β -sheet مركزية (اثنان منها على شكل متوازي والثلاث الأخرى متعاكسة) محاطة بقطعتين من نوع α -helice.



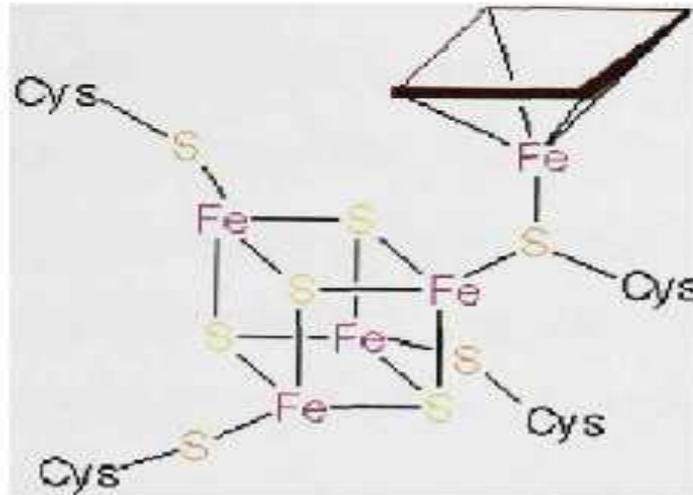
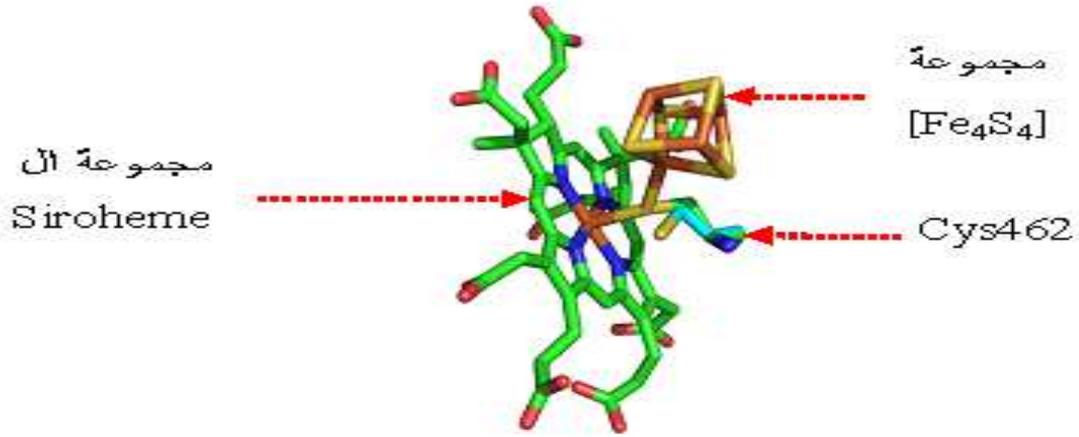
الشكل (21) : عرض البنية الثلاثية الأبعاد بالشكل الشريطي للمجال الثالث لـ Nitrite reductase بواسطة برنامج Pymol.



الشكل(22): يبين الموقع الفعال و تموضع المجموعتين غير البروتينيتين بشكل متقارب. العرض بواسطة برنامج Pymol.

3_ الجزء غير البروتيني لـ Nitrite reductase

يتضمن الـ Nitrite reductase مجموعتين غير بروتينيتين وهي مجموعة الـ Siroheme و مجموعة $[Fe_4S_4]$ ، هذان العاملان المساعدان يتوضعان على قرب وبشكل متقابل كما أنّ لهما رابطة مشتركة عن طريق الحمض الأميني Cys462، والذي يبعد بمسافة $2,31 \text{ \AA}$ عن ذرة الحديد المركزية لمجموعة الـ Siroheme، وبنفس المسافة عن ذرة الكبريت المقابلة له من مجموعة $[Fe_4S_4]$ كما هو موضح في (الشكل 23).

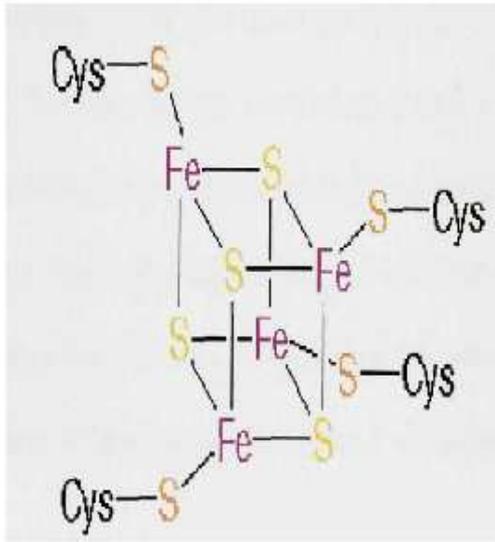


الشكل (23): يبين ارتباط مجموعة الـ Siroheme بمجموعة $[Fe_4S_4]$ على مستوى الموقع الفعال بواسطة الحمض الأميني Cys462 .

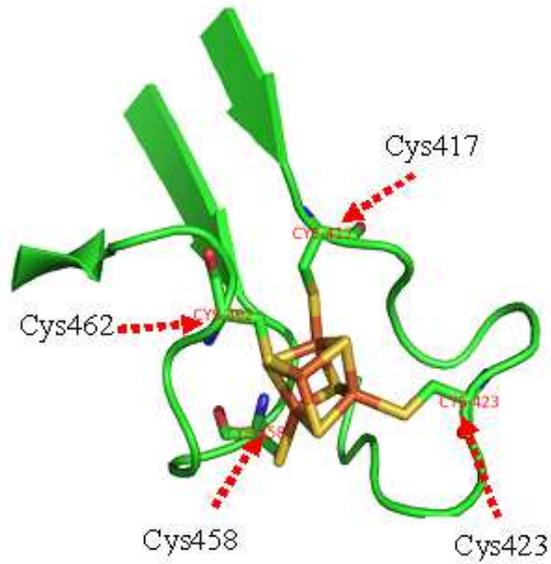
1.3_ مجموعة الـ [Fe₄S₄]

تحيط بمجموعة الـ [Fe₄S₄] عروتين (Loop). العروة الأولى مشكلة من سبعة أحماض أمينية تمتد من الحمض الأميني Cys417 حتى الحمض الأميني رقم Cys423 (CTGSQFC). أما العروة الثانية فتتكون من خمسة أحماض أمينية التي تمتد من الحمض الأميني Cys458 حتى الحمض الأميني رقم Cys462 (CPNTC). ترتبط مجموعة [Fe₄S₄] بأربع أحماض أمينية من نوع سيستيين، حيث ترتبط بالعروة الأولى عن طريق الحمض الأميني Cys417 والحمض الأميني Cys423 من جهة، وبالعروة الثانية بالحمض الأميني Cys458 و الحمض الأميني Cys462 من جهة أخرى كما هو موضح في (الشكل 24).

يكون هذا الارتباط عن طريق تشكيل روابط بين كل ذرة حديد لمجموعة الـ [Fe₄S₄] بذرات الكبريت لكل حمض أميني من الأحماض الأمينية Cys الأربعة السابقة وللتين تبعدان عن بعضهما بمسافة 2,31Å. الأحماض الأمينية السيستيين Cys الأربعة السابقة من المفترض أن تكون محفوظة عند البروتينات الأخرى التي تحتوي على مجموعة [Fe₄S₄] و التي لها الدور الأساسي في ربط وتثبيت مجموعة [Fe₄S₄].



ب

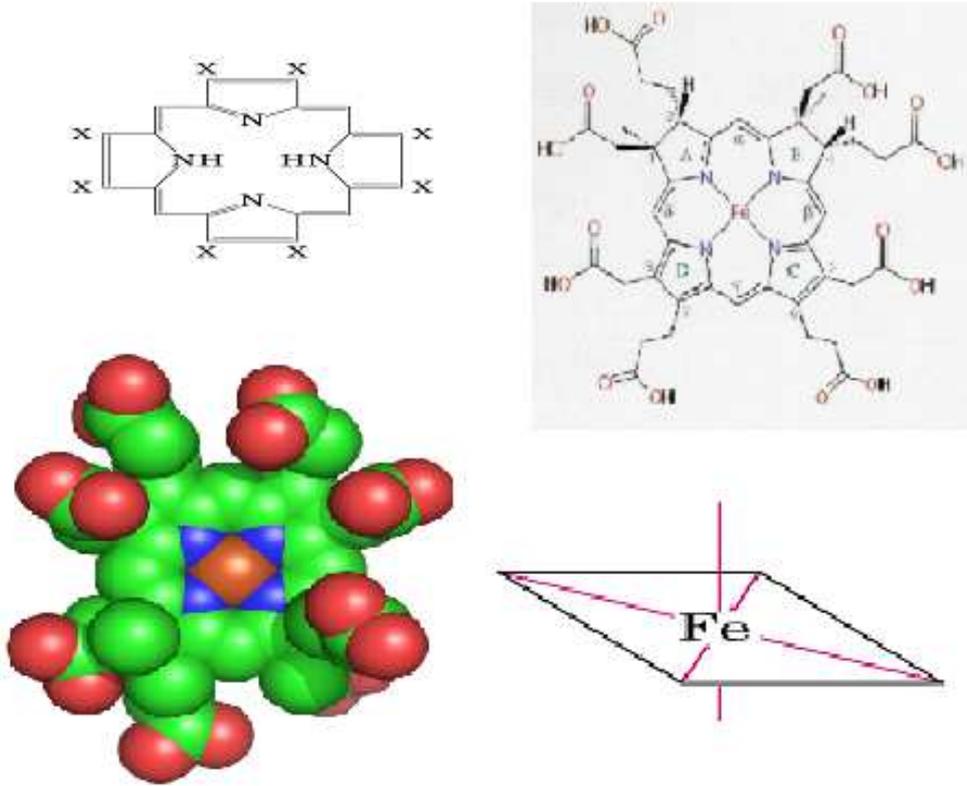


ا

الشكل (24): أ- عرض بواسطة برنامج Pymol لكيفية تثبيت مجموعة الـ [Fe₄S₄] بأربع أحماض أمينية (Cys417، Cys423، Cys458، Cys462). ب- الصيغة الفراغية للتثبيت.

2.3_ مجموعة الـ Siroheme

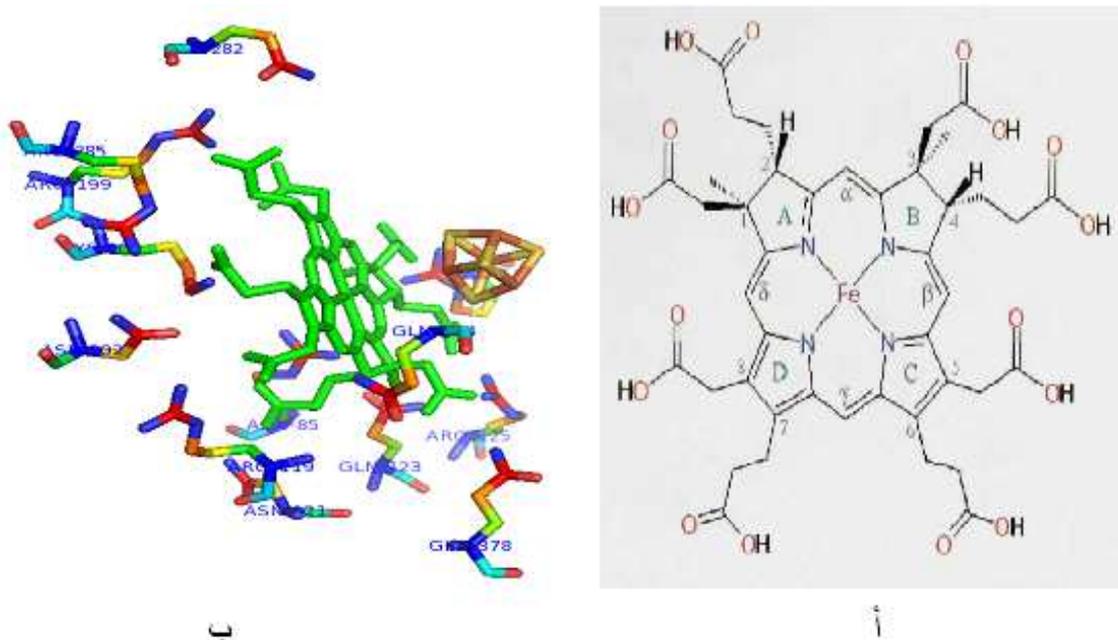
يتكون الهيم من جزء عضوي وذرة حديد، حيث يتكون الجزء العضوي Protoporphyrin من أربع حلقات من البيروول Pyrol متصلة ببعضها البعض لتشكل حلقة كبيرة واحدة تسمى حلقة رباعية البيروول. ترتبط ذرة الحديد الموجودة في مركز الـ Siroheme بذرات النتروجين الأربعة المحيطة بها و الموجودة في مركز حلقة Protoporphyrin ، كما يمكن لذرة الحديد أن تكون رابطين أخرتين على جانبي مستوى الهيم Heme التي تسمى بمواقع الارتباط التعاوني الخامسة والسادسة، حيث يمكن لذرة الحديد أن تأخذ حالتين من الأكسدة وتسمى حينئذ الحديدوز Ferrihem (2+) أو الحديدك Methem (3+).



الشكل (25): يبين نماذج مختلفة لعرض البنية الفراغية لمجموعة الـ Siroheme.

و من بين الصفات البارزة التي يتميز بها إنزيم الـ Nitrite reductase، هو وجود 7 أحماض أمينية متأينة بالقرب من مجموعة Siroheme وهي (Arg85. Arg74. Arg119. Arg200. Lys200. Arg199. Arg125. Arg285). والتي يفترض أنها المسؤولة عن استقرار وثبات

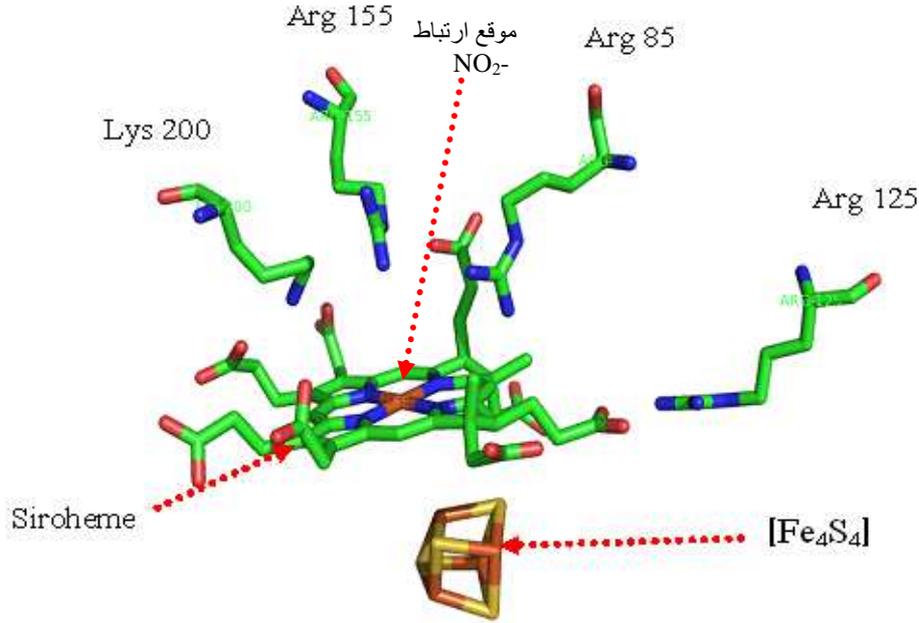
مجموعات الكربوكسيل لمجموعة الـ Siroheme، كما تساهم في ربط مادة التفاعل النترية، و في تفاعل تحويل وانتقال 6 الكترولونات إلى مادة التفاعل (Krueger et al., 1982).



الشكل (26): -أ- يبين الصيغة الفراغية ثنائية البعد لمجموعة الـ Siroheme . ب- يبين التوضع الفراغي ثلاثي الأبعاد لمجموعة الـ Siroheme محاطة بـ 7 أحماض أمينية متأينة، العرض بواسطة برنامج Pymol.

بينما الجانب الآخر لذرة الحديد (الجهة المعاكسة لتوضع مجموعة $[Fe_4S_4]$) من مجموعة الـ Siroheme لا يقرب أو يرتبط بها أي حمض أميني، بحيث أنّ اقرب ذرة لأقرب حمض أميني لا يقل بعدها عن ذرة الحديد بـ 5\AA . كما أنّ الأحماض الأمينية في هذه الجهة يشكلون جيبا مناسباً لارتباط مادة التفاعل (الشكل 27). وهذا ما يتوافق مع نتائج دراسات طيف الامتصاص Spectroscopic التي أشارت إلى أنّ النترية يرتبط بذرة الحديد من هذه الجهة (Kuznetsova et al., 2004). حيث تساهم أربع أحماض أمينية متأينة في ربط مادة التفاعل (النترية) بذرة الحديد المركزية لمجموعة الـ Siroheme وهي Arg155، Arg125، Arg85، و Lys200 وهي أحماض أمينية جد محفوظة عند إنزيم الـ Sulfite reductase أين يقابلها وجود الأحماض الأمينية المتأينة Arg83، His123، Arg 153، Lys 215 (Crane et al., 1995).

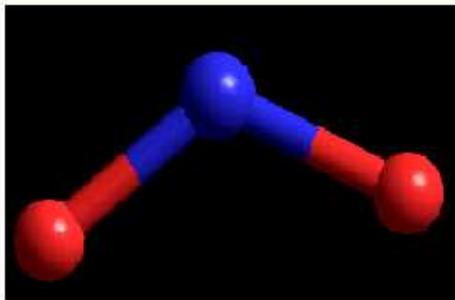
إن وجود العديد من الأحماض الأمينية المتأينة الموجبة في كلا الإنزيمين السابقين يلعب دوراً محورياً في وظيفة كل منهما، تتضمن ربط وتثبيت مادة التفاعل (SO_2^-) ذات الشحنة السالبة.



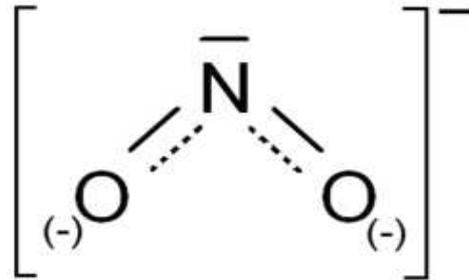
الشكل (27): يبين جهة ارتباط النتريت بال Siroheme (الموقع الفعال)، أين توجد أربع أحماض متأينة تساهم في هذا الربط Arg85 , Arg125 , Arg155, Lys200.

3.3_ مادة التفاعل النتريت

النتريت هي عبارة عن أملاح حمض النتروز (les sels de l'acide nitreux) الصيغة الكيميائية لها هي ON^{2-}



ب



أ

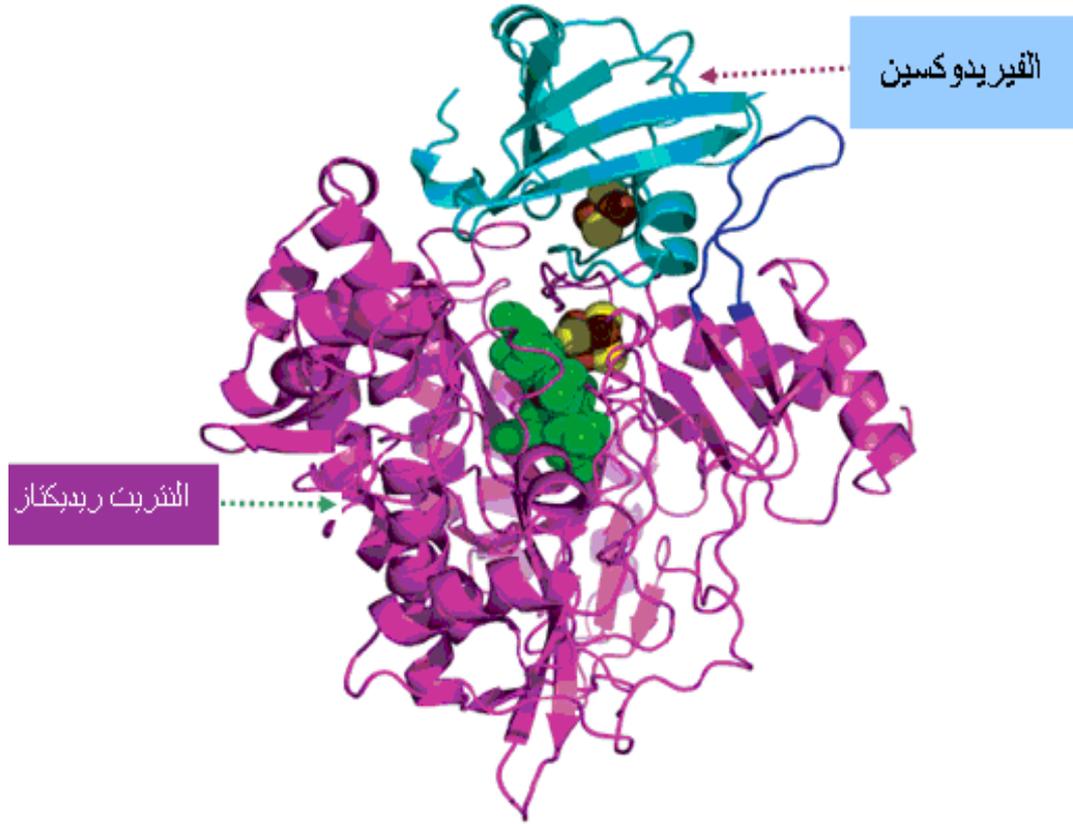
الشكل (28): أ- الصيغة الفراغية لأيون النتريت. ب- نموذج الكرة والعود لأيون النتريت

4_ آلية عمل إنزيم Nitrite reductase

تمتص جذور غالبية الباتات الراقية النتروجين في صورة نترات (NO_3^-) من التربة، ولكن النتروجين في صورته هذه لا يستعمل مباشرة من طرف النبات، إلا بعد أن تختزل النترات إلى الأمونيا NH_4^+ عبر مرحلتين أساسيتين قبل أن يدخل في تركيب المركبات النتروجينية المختلفة الموجودة في الخلايا النباتية.

تتضمن المرحلة الثانية اختزال مادة التفاعل النتريت (NO_2^-) إلى المركب النهائي الأمونيا إدماجه في تركيب المركبات النتروجينية NH_4^+ عن طريق إنزيم Nitrite reductase، هذا الأخير يستعمل الفيريدوكسين كمانح فيزيولوجي للإلكترونات (Knaff et al., 1991). يتواجد الـ Nitrite reductase في الصانعات الخضراء Chloroplasts بالنسبة للأنسجة اليخضورية، أين يُرجع الفيريدوكسين عن طريق النظام الضوئي الأول PSI. كما يتواجد الـ Nitrite reductase في الصانعات الملونة Proplasts بالنسبة للأنسجة غير اليخضورية أين يرجع الفيريدوكسين في هذه الحالة عن طريق NADPH (Knaff et al., 1991).

وجد أن عملية إرجاع النتريت إلى الأمونيا إلى 6 إلكترونات، حيث يحتوي إنزيم Nitrite reductase على موقع واحد للارتباط بالمرجع الفيزيولوجي الفيريدوكسين أين يرتبطا ارتباطا الكروستاتيكية ليشكلا مع بعض معقد (NIR-Fd) (Knaff et al., 1991)، أين يمكن للمجموعة غير البروتينية الوحيدة $[\text{Fe}_2\text{S}_2]$ للفيريدوكسين أن تمنح الكترونا واحدا في كل دفعة إلى المجموعة غير البروتينية $[\text{Fe}_4\text{S}_4]$ للـ Nitrite reductase أثناء دورة إرجاع الـ النتريت إلى الأمونيا حيث يستعمل الـ Nitrite reductase مجموعتيه غير البروتينيتين Siroheme و $[\text{Fe}_4\text{S}_4]$ لتجميع الإلكترونات الممنوحة من المرجع الفيزيولوجي الفيريدوكسين.



الشكل (29): يبين ارتباط ال Nitrite reductase (اللون البنفسجي) بالمرجع الطبيعي ال Ferredoxin (اللون الأزرق الفاتح).

دلت التجارب والدراسات السابقة أن Nitrite reductase في الحالة المعزولة تكون مجموعتيه غير البروتينيتين في حالة مؤكسدة كلياً $([Fe_4-S_4]^{2+}/Fe^{3+})$. حيث تكون ذرة الحديد لمجموعة Siroheme في حالة عزم عالي (High spin ferrec state). كما وجد أن ارتباط ال Nitrite reductase المؤكسد مع مادة التفاعل (NO_2^-) تقود الى تكوين المعقد المؤكسد $NIR-NO_2^-$ (Young et al., 1988).

ترتبط مادة التفاعل النتريت (NO_2^-) هذا الإنزيم في جيب أو شق عميق من البروتين وبالضبط بذرة الحديد المركزية لمجموعة ال Siroheme (Knaff et al., 1991; Kuznetsova et al., 2004). ولقد أظهر التحليل الطيفي المغناطيسي EPR أن ذرة الحديد المركزية لمجموعة ال Siroheme يمكن أن تحتوي على خمسة روابط أو ستة روابط حيث تكون الرابطة السادسة

مركزية وعمودية على مجموعة ال Siroheme وهي رابطة ضعيفة ترتبط بمادة التفاعل النترت في حالة وجودها.

ترتبط ذرة الحديد بأربعة روابط مع ذرات النتروجين الأربعة الموجودة في مركز حلقة Protoporphyrin كما تستطيع ذرة الحديد أن تكون رابطتين أخرتين على جانبي مستوى الهيم التي تسمى بمواقع الارتباط التعاوني الخامسة والسادسة، حيث يمكن لذرة الحديد أن تأخذ حالتين من الأكسدة، الحالة الأولى تكون بالشكل Fe^{2+} وتسمى حينئذ الحديدوز Ferrihem $(2+)$ أو الحالة الثانية أين تكون بالشكل Fe^{3+} وتسمى عندئذ الحديدك Methem $(3+)$ ، الرابطة الخامسة التي تكون من جهة مجموعة $[Fe_4S_4]$ ترتبط بالحمض الاميني Cys462 هذا الأخير يرتبط أيضا بمجموعة $[Fe_4S_4]$ بواسطة جسر كبريتي، في حالة غياب مادة التفاعل من الممكن أن ترتبط الرابطة السادسة بشكل ضعيف بجزء الماء (Kuznetsova et al., 2004).

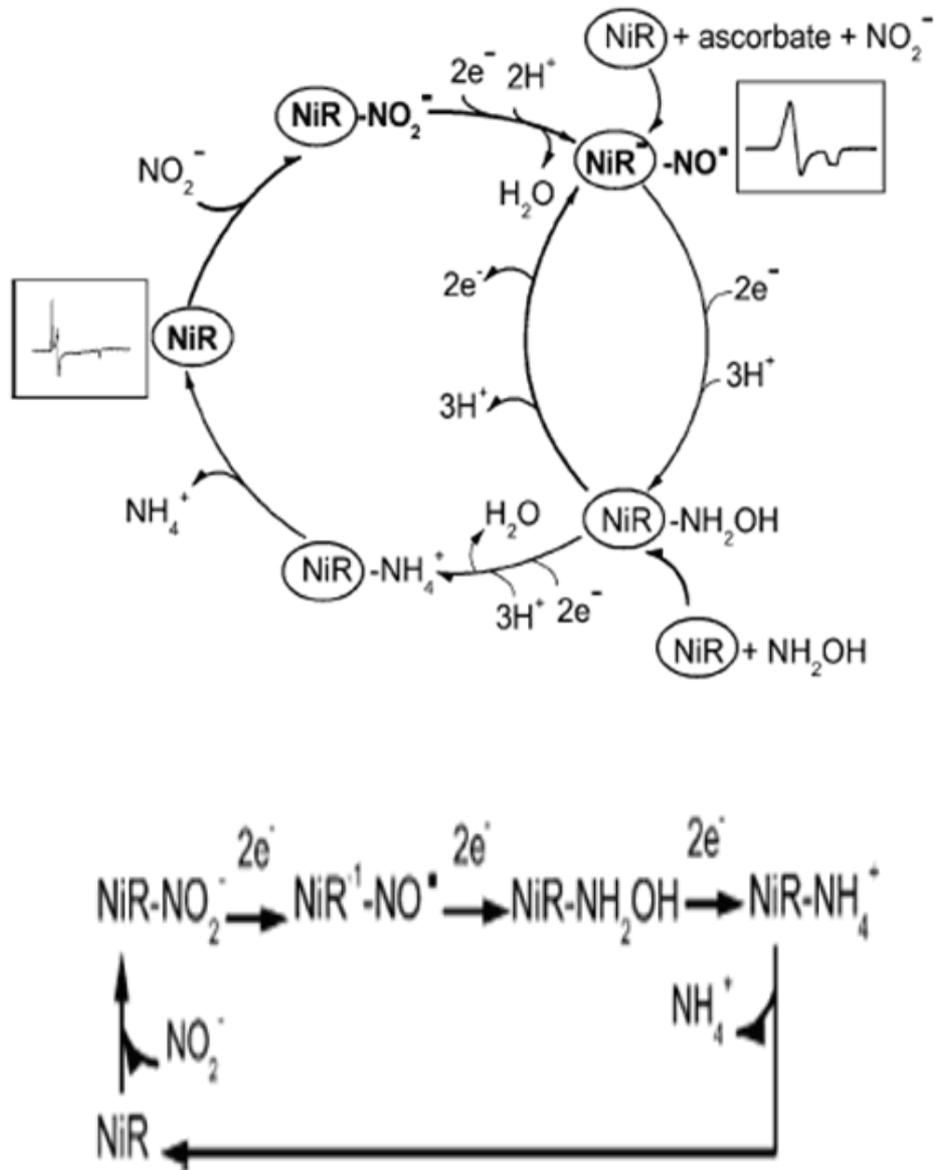
إنّ استقرار ارتباط النترت (NO_2^-) بذرة الحديد يضمنه وجود العديد من الأحماض الأمينية القاعدية كال Arg و Lys ذات الشحنة الموجبة والتي تجذب ايون النترت ذو الشحنة السالبة. وعلى الرغم من ارتباط وكذا المسافة القريبة التي تجمع بين مجموعة ال Siroheme ومجموعة ال $[Fe_4S_4]$ إلا أنّ عمل كل منهما مستقل يكون مستقلا. حيث تستقبل مجموعة $[Fe_4S_4]$ الالكترونات من المرجع الطبيعي ال Ferredoxin وتحولها نحو مجموعة Siroheme هذه الأخيرة تحولها بدورها إلى مادة التفاعل النترت التي تبقى مرتبطة حتى إرجاعها إلى المركب النهائي وهو الأمونيا NH_4^+ (Kuznetsova et al., 2004).

إن عملية تحويل مادة التفاعل النترت (NO_2^-) إلى أمونيا (NH_4^+) تتطلب على الأقل خمسة مراحل: حيث تشير الأدلة بوضوح على أنه في جميع الاحتمالات فإن الإنزيم يرجع في الأول بواسطة إلكترون واحد إلى الحالة التي تصبح فيها ذرة الحديد لمجموعة ال Siroheme بالشكل Fe^{2+} والتي ترتبط بسرعة بالنترت (بالنسبة للمجموعة $[Fe_4S_4]^{2+}$ ترجع مؤقتا إلى الحالة $[Fe_4S_4]^{1+}$ قبل أن تمنح الإلكترون إلى مجموعة Siroheme لترجع إلى الحالة الأصلية $^{2+}$ $[Fe_4S_4]$).

بعد ذلك تستقبل مجموعة $[Fe_4S_4]$ إلكترون آخر والذي ستحوله نحو مجموعة ال Siroheme ليشكل المعقد (NIR-NO) Ferrosiroheme. وصول إلكترونين آخرين إلى المعقد

يمكن أن يرجع هذا الأخير إلى المركب الوسيطى الـ Hydroxylamine (NH₂OH) مع أنه لم يتم تحديده لحد الآن (Uma et al., 2005).

وأخيرا فإن وصول إلكترونين آخرين إلى المعقد NIR-NH₂OH سيؤدي إلى تشكيل المركب النهائى أي الأمونى NH₄⁺ (Kuznetsova et al., 2004).



الشكل (30): يبين مراحل إرجاع النترت إلى أمونيا بواسطة Nitrite reductase (Kuznetsova et al., 2004).

الخلاصة

تسمح طريقة النمذجة بالتماثل Homology modeling وباستعمال أدوات وبرامج البيومعلوماتية Bioinformatic tools ببناء نماذج نظرية ثلاثية الأبعاد لبنية البروتين، هذه الأخيرة تعتبر المفتاح الأساسي لفهم آلية عمل البروتينات. حيث ظهرت النمذجة بالتماثل كبديل لتعويض عجز الطرق التجريبية (X-ray. RMN) في تحديد أعداد كافية من البنيات البروتينية.

قمنا في هذا الإطار وباستعمال طريقة النمذجة بالتماثل وبرنامج Modeller بصفة أساسية ببناء عشرة نماذج نظرية ثلاثية الأبعاد لإنزيم Ferredoxin-nitrite reductase لنبات *Arabidopsis thaliana* ، وذلك عبر أربع خطوات رئيسية. تم اختيار النموذج الذي يملك الاستقرار الطاقوي الأكبر من بينها.

دلت كل نتائج المعايير المعتمدة في تقييم النموذج المتحصل عليه على انه يملك بنية ذات خصائص جيدة تقربه من البنية المحددة تجريبيا. حيث تحصلنا على نسبة تشابه معتبرة بين تتابع البروتين الهدف والبروتين المعتمد كقالب وصلت إلى حوالي 81% وهي نسبة تسمح باعتماد طريقة النمذجة بالتماثل والحصول على بنيات جيدة، والتي أدت بدورها الى الحصول على نسبة 79% كمطابقة بين تتابع البروتين الهدف وبنية القالب وهي نسبة عالية تعبر عن جودة النماذج المتحصل عليها.

أظهرت نتائج التوضع الفراغي للأحماض الأمينية أن نسبة 99,5% منها سليم وينتمي للمناطق المسموح بها، نتائج قياس الاستقرار الطاقوي أظهرت أيضا أن الحالة الطاقوية لبنية النموذج المتحصل عليه قريبة جدا من الحالة الطاقوية لبنية البروتين القالب المحددة تجريبيا. وفيما يتعلق بدرجة التطابق بين هيكلي سلسلي بنيتي البروتين النموذج والقالب فكانت قيمة RMSD بينهما تقارب 0,2Å ، كما أن قيمة Z scores لبنية النموذج تساوي -10.37.

تدل النتائج السابقة على مدى نجاح عملية النمذجة بالتماثل في بناء نموذج لبنية نظرية ثلاثية الأبعاد لإنزيم Ferredoxin-nitrite reductase لنبات *Arabidopsis thaliana* ذات مواصفات قريبة جدا من البنيات المحددة تجريبيا والتي يمكن اعتمادها في الدراسة البنوية و الوظيفية له.

أظهرت الدراسة التحليلية لبنية النموذج أنها بنية ثلثية تتكون من تتابع 509 حمض أميني يمكن تقسيمها إلى ثلاث مجالات رئيسية، كما تحتوي على شق عميق به مجموعتين غير بروتينيتين (Siroheme و [Fe₄S₄]) والذي يعتبر الموقع الفعال لهذا الإنزيم.

وبالاعتماد على الدراسات السابقة تم اقتراح آلية عمل لهذا الإنزيم في تحويل مادة التفاعل النترت (NO₂⁻) إلى أمونيا (NH₄⁺) حيث تتطلب على الأقل خمسة مراحل: في البداية يُرجَّع الإنزيم بواسطة إلكترون واحد إلى الحالة التي تصبح فيها ذرة الحديد لمجموعة Siroheme بالشكل Fe₂⁺ التي ترتبط بسرعة بالنترت، بعد ذلك تستقبل مجموعة [Fe₄S₄] إلكترون آخر والذي ستحوّله نحو مجموعة ال Siroheme ليتشكل المعقد (NIR-NO) NO/Ferrosiroheme، وصول إلكترونين آخرين إلى المعقد يمكن أن تُرجَّع هذا الأخير إلى المركب الوسطي Hydroxylamine (NH₂OH)، وأخيرا فان وصول إلكترونين إضافيين إلى المعقد NIR-NH₂OH سيؤدي إلى تشكيل المركب النهائي الأمونيا NH₄⁺.

آفاق وتطلعات

يعتبر هذا العمل كبدائية لدراسة وفهم بنيات وآليات عمل بعض الإنزيمات المهمة في الوظائف الخلوية النباتية في الظروف البيئية غير الملائمة كالإجهاد المائي، الإجهاد الحراري، وجود المثبطات. التي نطمح لدراستها مستقبلا بواسطة دمج الطرق التجريبية المخبرية و الطرق البيومعلوماتية.

المراجع

Abola E.E., Bernstein F.C., Bryant S.H., Koetzle T.F., Weng J., 1987, Protein Data Bank in Crystallographic Databases-Information Content, Software Systems, Scientific Applications, eds. F.H. Allen, G. Bergerhoff, and R. Sievers, Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester, pp. 107-132.

Alberts B., Bray D., Lewis J., Raff, M., Toberts K., Watson J., 1994, Molecular Biology of the Cell. Garland Publishing, Inc. New York, London.

Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J., 1990, Basic local alignment search tool, *J. Mol. Biol.*, **215**:403–10.

Altschul S.F., Boguski M.S., Gish W., Wootton J.C., 1994, Issues in searching molecular sequence databases, *Nat. Genet.*, **6**:119–29.

Altschul S.F., Madden T.L., Schaffer A.A, Zhang J.Z., Miller W., Lipman D.J., 1997, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.*, **25**:3389–402.

Anfinsen C.B., 1973, Structure of a protein is determined solely by the amino acids sequence info, *Science*, **181**:223–230.

Anfinsen C.B., Flavin M., Farnsworth J., 1952, Preliminary studies on ribonuclease structure, *Biochim Biophys Acta*, **9**:468-9.

Anfinsen C. B., Redfield R. R., Choate W. L., Page J., Carroll W. R., 1954, Studies on the gross structure, cross-linkages, and terminal sequences in ribonuclease, *The Journal of Biological Chemistry*, **207**:201–210.

Anfinsen C.B., 1973, Principles that govern the folding of protein chains, *Science* **181**: 223-230.

Anfinsen C.B., Haber E., Sela M., White J.r., 1961, The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain, *Proc Natl Acad Sci U S A* **47** :1309–1314.

Aparicio P., Knaff J., Malkin R., 1975, Arch. Biochem, *Biophys*, **169**: 102-107.

Apostolico A., Giancarlo R., 1998, Sequence alignment in molecular biology, *J. Comput. Biol.*, **5**:173–96.

Archibald J.M., Rogers M.B., Troop M., Ishida K., Keeling P.J., 2003, *Proc Natl Acad Sci U S A*, **100**: 7678-7683.

Arnon D.I., 1988 , The discovery of ferredoxin: the photosynthetic path. *Trends Biochem Sci. Jan*, **13**:30–33.

- Aszódi A., Taylor W.R.**, 1996, Homology modelling by distance geometry, *Folding & Design (now Structure)* **1**: 325-334.
- Back E., Burkhardt W., Moyer M., Privalle L., Rothstein S.**, 1988, *Mol. Gen. Genet*, **212**: 20-26.
- Bairoch A., Apweiler R.**, 1999, The SWISSPROT protein sequence data bank and its supplement TrEMBL in 1999, *Nucleic Acids Res*, **27**:49–54.
- Bairoch A., Apweiler R.**, 2000, The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000, *Nucleic Acids Res*, **28**: 45–48.
- Bajorath J., Stenkamp R., Aruffo A.**, 1994, Knowledge-based model building of proteins: concepts and examples, *Protein Sci*, **2**:1798–1810.
- Baker D., Sali A.**, 2001, Protein structure prediction and structuralgenomics, *Science*, **294**:93-96.
- Barton G.J., Sternberg M.J.**, 1987, A strategy for the rapid multiple alignment of protein sequences; confidence levels from tertiary structure comparisons, *J. Mol. Biol*, **198**:327–37.
- Barton G.J.**, 1998, Protein sequence alignment and database scanning. In Protein Structure Prediction: A Practical Approach, ed. MJE Sternberg. Oxford, UK: *Oxford Univ Press*
- Baxevanis A.D.**, 1998, Practical aspects of multiple sequence alignment, *Methods Biochem Anal*, **39**:172–88.
- Bellissimo D.B., Privalle, L. S.**, 1995, *Arch Biochem Biophys* **323**:155-163.
- Benson D.A., Karsch I., Lipman D.J., Ostell J., Rapp B.A., Wheeler D.L.**, 2000, Gen-Bank, *Nucleic Acid Res*, **28**:15–18.
- Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E.**, 2000, The Protein Data Bank, *Nucleic Acids Res*, **28**:235–42.
- Berman H.M., Henrick K., Nakamura H.**, 2003, Announcing the worldwide Protein Data Bank. *Nature Structural Biology* **10** : p. 980.
- Berman H., Henrick K., Nakamura H., Markley J.L.**, 2006, The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res*.**35**:D301-3.
- Bernal A., Ear U., Kyripides N.**, 2001, Genomes OnLine Database (GOLD): a monitor of genome projects world-wide, *Nucl Acids Res*, **29**: 126-127
- Bernstein F.C., Koetzle T.F., Williams G.J., Meyer E.F., Brice M.D., Rodgers J.R., Kennard O., Shimanouchi T., Tasumi M.**, 1977. The protein data bank: a computer-based archival file for macromolecular structures, *J. Mol. Biol*, **112**:535–42.

Bhattacharya A., Tejero R., Montelione G.T., 2007, Evaluating Protein Structures Determined by Structural Genomics Consortia, *Proteins*, **66**:778-95.

Bino J., Andrej S., 2003 Comparative protein structure modeling by iterative alignment, model building and model assessment. *Nucl. Acids Res.* **31**: 3982-3992.

Blundell T.L., Sibanda B.L., Sternberg M.J., Thornton J.M., 1987, Knowledge-based prediction of protein structures and the design of novel molecules, *Nature*, **326**:347–52.

Boeckmann B., Bairoch A., Apweiler R., Blatter M.C., Estreicher A., Gasteiger E., Martin M.J., Michoud K., O'Donovan C., Phan I., 2003, The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003, *Nucl Acids Res*, **31**: 365-370.

Bowie J.U., Luthy R., Eisenberg D., 1991, A method to identify protein sequences that fold into a known three-dimensional structure, *Science* **253**:164-70

Bowsher C.G., Emes M.J., Cammack R., Hucklesby D.P., 1988, Purification and properties of nitrite reductase from roots of pea (*Pisum sativum* cv. Meteor). *Planta* **175**: 334-340.

Bowsher C.G., Hucklesby D.P., Emes M.J., 1989, Nitrite reduction and carbohydrate metabolism in plastids purified from roots of *Pisum sativum* L. *Planta*, **177**: 359-366.

Bowsher C.G., Boulton E.L., Rose J., Mayagam S., Emes M.J., 1992, Reductant for glutamate synthase in generated by the oxidative pentose phosphate pathway in non-photosynthetic root plastids . *Plant J*, **2**: 893-898.

Brenner S.E., Chothia C., Hubbard T.J., 1998, Assessing sequence comparison methods with reliable structurally identified distant evolutionary relationships, *Proc Natl Acad Sci U S A*, **95**:6073-6078.

Briffeuil P., Baudoux G., Lambert C., Bolle X., Vinals C., Feytmans E., Depiereux E., 1998, Comparative analysis of seven multiple protein sequence alignment servers: clues to enhance reliability of predictions, *Bioinformatics*, **14**:357–66.

Brocklehurst S.M., Perham R.N., 1993, Prediction of the three-dimensional structures of the biotinylated domain from yeast pyruvate carboxylase and of the lipoylated H-protein from the pea leaf glycine cleavage system: a new automated method for the prediction of protein tertiary structure, *Prot. Sci*, **2**: 626–639.

Campbell W.H., Kinghorn J.R., 1990, Functional domains of assimilatory nitrate reductases and nitrite reductases. *Trends Biochem. Sci.* **15**: 315-319.

Cammack R., Hucklesby D.P., Hewitt E.J., 1978, EPR studies on the mechanism of leaf nitrite reductase. Signals from the iron-sulphur centres and haem under turnover conditions. *Biochem. J.* **171**, 519-526.

Cammack R., 1984, in Charge and Field Effects in Biosystems (Allen, M. J., and Usherwood, P. N. R., eds) pp. 41–51, *Abacus Press*, Tunbridge Wells.

Cammack R., Rao K. K., Hall D.O., 1985, Ferredoxins: structure and function of a ubiquitous group of proteins. *Physiol. Veg.* **23**: 649-658.

Chothia C., Lesk A.M., 1986, The relation between the divergence of sequence and structure in proteins, *EMBO J*, **5**:823–26.

Christner J.A., Münck E., Janick P.A. Siegel W.M., 1981, Mössbauer Spectroscopic Studies of Escherichia coli Sulfite Reductase, *J. Biol. Chem.* **256**: 2098-2101.

Claessens M., Van Cutsem E., Lasters I., Wodak S., 1989, Modelling the polypeptide backbone with 'spare parts' from known protein structures, *Protein Eng.* **2**:335-345.

Crane B.R., Siegel L.M., Getzoff E.D., 1995, Sulfite reductase structure at 1.6 Å: evolution and catalysis for reduction of inorganic anions. *Science*, **270**:59-67.

Crane B.R., Siegel L.M., Getzoff E.D., 1997, Structures of the siroheme- and Fe4S4-containing active center of sulfite reductase in different states of oxidation: heme activation via reduction-gated exogenous ligand exchange. *Biochemistry*, **36**:12101-19.

Crawford N.M., Glass D.M., 1998, Molecular and physiological aspects of nitrate uptake in plants. *Trends Plant Sci*, **3**: 389–395.

Daniel I. A., George M.S., 1988, Cytochrome b-559 and proton conductance in oxygenic photosynthesis. *Biophysics*, **85**: 9524-9528.

Dai S., Schwendtmayer C., Schurmann P., Ramaswamy S., Eklund H., 2000: Redox signaling in chloroplasts: cleavage of disulfides by an iron-sulfur cluster. *Science*, **287**:655-658.

David L.Wheeler., Tanya Barrett., Dennis A. Benson, Stephen H. Bryant, Kathi Canese, Vyacheslav Chetvernin, Deanna M. Church, Michael DiCuccio, Ron Edgar, Scott Federhen, Lewis Y. Geer, Yuri Kapustin, Oleg Khovayko, David Landsman, David J. Lipman, Thomas L. Madden, Donna R. Maglott, James Ostell, Vadim Miller, Kim D. Pruitt, Gregory D. Schuler, Edwin Sequeira, Steven T. Sherry, Karl Sirotkin, Alexandre Souvorov, Grigory Starchenko, Roman L. Tatusov, Tatiana A. Tatusova, Lukas Wagner, Eugene Yaschenko., 2007, Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research* **35**: 5-12.

Day E.P., Peterson J., Bonvoisin J.J., Young L.J., Wilkerson J.O., Siegel L.M., 1988, Magnetization of the sulfite and nitrite complexes of oxidized sulfite and nitrite reductases: EPR silent spin $S = 1/2$ states. *Biochemistry*, **27**: 2126-2132.

Dinner A. R., Sali A., Karplus M., 1996, The folding mechanism of larger model proteins: Role of native structure. *Proceedings of the National Academy of Sciences of the United States of America*, **93**:8356–8361

Dobson C.M., 2004, Principles of protein folding, misfolding and aggregation, *Seminars in Cell & Developmental Biology*, **15**: 3-16.

Dose M.M., Hirasawa M., Kleis-SanFrancisco S., Lew E.L., Knaff D.B.,1997, The ferredoxin-binding site of ferredoxin:nitrite oxidoreductase. Differential chemical modification of the free enzyme and its complex with ferredoxin. *Plant Physiol.* **114**: 1047-1053.

Eisenberg D., 2003, The discovery of the α -helix and β -sheet, the principal structural features of proteins, *Proc Natl Acad Sci USA*, **100**: 11207-11210.

Eswar N., Madhusudhan M.S., Marti-Renom M.A., Sali A., 2005, BUILD_PROFILE: a module for calculating sequence profiles in MODELLER <http://www.salilab.org/modeller>

Felsenstein J., 1985, Confidence limits on phylogenies: an approach using the bootstrap, *Evolution*, **39**:783.

Fitz-Gibbon S., Ladner H., Kim U.J., Stetter K.O., Simon M.I., Miller J.H., 2002, Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum*. *Proc. Natl. Acad. Sci. U.S.A.* **99**: 984-989.

Fiser A., 2000, Modeling of loops in protein structures, *Protein Sci*, **9**:1753–1773.

Fischer D., Eisenberg D., 1997, Assigning folds to the proteins encoded by the genome of *Mycoplasma genitalium*, *Proc Natl Acad Sci USA*, **94**:11929–34.

Foust G.P., Mayhew S.G., Massey V., 1969, Complex Formation between Ferredoxin Triphosphopyridine Nucleotide Reductase and Electron Transfer Proteins. *J. Biol. Chem.* **244**: 964-970.

Fowden, L., 1979, Nitrogen Assimilation in Plants, *Academic Press*, New York, 166,557.

Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M., Fritchman J.L., Weidman J.F., Small Krueger R.J., Siegel L.M., 1982, Evidence for siroheme-Fe₄S₄ interaction in spinach ferredoxin-sulfite reductase. *Biochemistry.* **8**:2905–2909.

Friemann A., Brinkmann K., Hachtel W., 1992, Sequence of a cDNA encoding nitrite reductase from the tree *Betula pendula* and identification of conserved protein regions. *Mol Gen Genet*, **231**: 411–416.

Fry I.V., Cammack R., Hucklesby D.P., Hewitt E.J., 1980, Stability of the nitrosyl-sirohaem complex of plant nitrite reductase, investigated by EPR spectroscopy. *FEBS Lett.* **111**: 377-380.

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A., 2005, Protein Identification and Analysis Tools on the ExPASy Server, in *The Proteomics Protocols Handbook*, Walker, J.M., Editor., Humana Press. p. 571-607.

Glöckner F.O., Kube M., Bauer M., Teeling H., Lombardot T., Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R., Schlesner H., Amann R. Reinhardt R., 2003 Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. *Proc. Natl. Acad. Sci. USA*, **100** :8298-8303.

Goldman B.S., Roth J.R., 1993, Genetic structure and regulation of the *cysG* gene in *Salmonella typhimurium*. *J Bacteriol*, **175**: 1457-1466.

Greer J., 1990, Comparative modeling methods: application to the family of the mammalian serine proteases, *Proteins*, **7**:317-334.

Gribskov M., Veretnik S., 1996, Identification of sequence pattern with profile analysis, *Methods Enzymol*, **266**:198-212.

Haas B. J., 2003, The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA.

Hilliard N.P., Hirasawa M., Knaff D.B., Shaw R.W., 1991, A reexamination of the properties of spinach nitrite reductase: protein and siroheme content heterogeneity in purified preparations. *Arch. Biochem.Biophys.* **291**: 195-199.

Hirasawa M., Tamura G., 1980, Ferredoxin-dependent nitrite reductase from spinach leaves. *Agric. Biol. Chem.* **44**: 749-758.

Hirasawa-Soga M., Tamura G., 1981, Some properties of ferredoxin-nitrite reductase from *Spinacia oleracea*. *Agric. Biol. Chem*, **45**: 1615-1620.

Hirasawa M., Knaff D.B., 1985, Interaction of ferredoxin-linked nitrite reductase with ferredoxin. *Biochim. Biophys. Acta*, **830**: 173-180.

Hirasawa M., Boyer J.M., Gray K.A., Davis D.J., Knaff D.B., 1986, The interaction of ferredoxin with ferredoxin-linked enzymes. *Bio- chim Biophys Acta*, **851**: 23-28.

Hirasawa M., Boyer J.M., Gray K.A., Davis D.J., Knaff D.B., 1987, The interaction of ferredoxin-linked sulfite reductase with ferredoxin. *FEBS Lett*, **221**:343-348.

Hirasawa M., Droux M., Gray K.A., Boyer J.M., Davis D.J., Buchanan B.B., Knaff D.B., 1988, Ferredoxin-thioredoxin reductase: Properties of its complex with ferredoxin, *Biochim Biophys Acta*, **935**: 1-8.

Hirasawa M., Shaw R.W., Palmer G., Knaff D.B., 1987, Prosthetic group content and ligand-binding properties of spinach nitrite reductase. *J. Biol. Chem.* **262**:12428-12433.

- Hirasawa M., Gray K.A., Sung J.D., Knaff D.B.**, 1989, Spinach nitrite reductase: subunit composition and siroheme redox potential. *Arch Biochem Biophys.* **275**: 1-10.
- Hucklesby D.P., James D.M., Banwell M.J., Hewitt E.J.**, 1976, Properties of nitrite reductase from Cucurbita pepo. *Phytochem.* **15**: 599-603.
- Humphrey W., Dalke A., Schulten K.**, 1996, VMD – Visual Molecular Dynamics. *J Mol Graph*, **14**:33-38.
- Hirasawa M., Tollin G., Salamon Z., Knaff D.B.**, 1994, Transient kinetic and oxidation-reduction studies of spinach ferredoxin: nitrite oxidoreductase. *Biochim Biophys Acta*, **1185**: 336-345.
- Hirasawa M., Rubio L.M., Griffin J.L., Flores E., Herrero A., Li J., Kim S.K., Hurley J.K., Tollin G., Knaff D.B.**, 2004, Complex formation between ferredoxin and Synechococcus ferredoxin : nitrate oxidoreductase. *Biochim. Biophys. Acta*, **1608**: 155-162.
- Hirsch R.E., Sussman M.R.**, 1999, Improving nutrient capture from soil by the genetic manipulation of crop plants. *Trends Biotechnol*, **17**: 356-361.
- Henikoff J.G., Pietrokovski S., McCallum C.M., Henikoff S.**, 2000, Blocks-based methods for detecting protein homology, *Electrophoresis*, **21**:1700-1706.
- Holm L., Sander C.**, 1996, Alignment of three-dimensional protein structures : network server for database searching, *Methods Enzymol*, **266** : 653–62.
- Holm L., Sander C.**, 1996, Mapping the protein universe. *Science*, **273**:595–602.
- Holm L., Sander C.**, 1999, Protein folds and families: sequence and structure alignments, *Nucleic Acids Res*, **27**:244–47.
- Hooft R.W., Vriend G., Sander C., Abola E.E.**, 1996, Errors in protein structures, *Nature*, 381:272.
- Hubbard T.J., Ailey B., Brenner S.E., Murzin A.G., Chothia C.**, 1999, SCOP: A structural classification of proteins database, *Nucleic Acids Res*, **27**:254–56.
- Huynen M., Doerks T., Eisenhaber F., Orengo C., Sunyaev S., Yuan Y., Bork P.** , 1998, Homology-based fold predictions for Mycoplasma genitalium proteins, *J Mol Biol*, **280**:323–26.
- Ida S., Mikami B.**, 1986, a purification procedure and characterization of chemical properties, *Biochim. Biophys. Acta*, **871**: 167-176.
- Ip S.M., Kerr J., IngledeW W.J., Wray J.L.**, 1990, Purification and characterisation of barley leaf nitrite reductase. *Plant Sci.* **66**:155-165.
- Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B., Kapatral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A., Chu L., Mazur M.**,

- Goltsman E., Larsen N., D'Souza M., Walunas T., Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Dusko Ehrlich S., Overbeek R., Kyrpides N.,** 2003, Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. *Nature*. **423**:87-91.
- Janick P.A., Siegel L.M.,** 1982, Electron paramagnetic resonance and optical spectroscopic evidence for interaction between siroheme and Fe4S4 prosthetic groups in *Escherichia coli* sulfite reductase hemoprotein subunit. *Biochemistry*, **21**: 3538-3547.
- Jansen R., Gerstein M.,** 2000, Analysis of the yeast transcriptome with structural and functional categories: characterizing highly expressed proteins, *Nucl acids Res*, **28**:1481-8.
- Jeanmougin F., Thompson J.D., Gouy M., Gibson D.G., Higgins T.J.,** 1998, Multiple sequence alignment with CLUSTAL X. *Trends Biochem, Sci*, **23**:403-5.
- Johnson M.S., Overington J.P.,** 1993, A structural basis for sequence comparisons: an evaluation of scoring methodologies, *J Mol Biol*, **233**:716-38.
- Johnson M.S., Srinivasan N., Sowdhamini R., Blundell T.L.,** 1994, Knowledge-based protein modelling, *Mol Biol*, **29**:1-68.
- Jones D.T.,** 1999, GenTHREADER: An efficient and reliable protein fold recognition method for genomic sequences, *J Mol Biol*, **287**:797-815.
- Jones, D. T.,** 2000. A practical guide to protein structure prediction. *Methods in Molecular Biology*, **143**:131-154.
- Jones T.A., Thirup S.,** 1986, Using known substructures in protein model building and crystallography, *EMBO J*, **5**:819-822.
- Kaneko T.S., Sato, H. Kotani A., Tanaka E., Asamizu Y., Nakamura N., Miyajima M., Hirosawa M., Sugiura S., Sasamoto T., Kimura T., Hosouchi A., Matsuno A., Muraki N., Nakazaki K., Naruo S., Okumura S., Shinpo C., Takeuchi T., Wada A., Watanabe M., Yamada M.,** 1996, Sequence analysis of the genome of the unicellular *Cyanobacterium Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res*. **3**:109-136.
- Kendrew J.C., Bodo G., Dintzis H.M., Parrish R.G., Wyckoff H., Phillips D.C.,** 1958, A three-dimensional model of the myoglobin molecule obtained by x-ray analysis, *Nature* **181**: 662-666.
- Kim CH., Cho Y.H., Hong Y.N.,** 1999, Regulation of nitrite reductase by light and nitrate in the cotyledons of hot pepper (*Capsicum annuum L.*). *Mol Cells*.**9**:152-157.
- Knaff D.B., Hirasawa M.,** 1991, Ferredoxin-dependent chloroplast enzymes, *Biochim Biophys Acta*, **1056**: 93-125.

- Knaff D. B.**, 1996, Oxygenic Photosynthesis; The Light Reactions, *Kluwer Academic, Dordrecht*, The Netherlands p. 333-361.
- Koehl P., Levitt M.**, 1999, A brighter future for protein structure prediction, *Nat Struct Biol*, **6**:108–11.
- Kopp J., Schwede T.**, 2004, Automated protein structure homology modeling: a progress report, *Pharmacogenomics*, **5**:405–16.
- Kronenberger J., Lepingle A., Caboche M., Vaucheret H.**, 1993, Cloning and expression of distinct nitrite reductases in tobacco leaves and roots. *Mol Gen Genet*. **236**:203–208.
- Krueger R.J., Siegel L.M.**, 1982, Spinach siroheme enzymes: Isolation and characterization of ferredoxin-sulfite reductase and comparison of properties with ferredoxin-nitrite reductase. *Biochemistry*. **8**:2892–2904.
- Kuznetsova S., Knaff D.B., Hirasawa M., Lagoutte B., Setif P.**, 2004 , Mechanism of spinach chloroplast ferredoxin-dependent nitrite reductase: Spectroscopic evidence for intermediate states, *Biochemistry* **43**:510-517.
- Kuznetsova S., Knaff D.B., Hirasawa M., Setif P., Mattioli T.A.**, 2004, Reactions of spinach nitrite reductase with its substrate, nitrite, and a putative intermediate hydroxylamine, *Biochemistry*, **43**:10765-10774.
- Lahners K., Kramer V., Back E., Privalle L., Rothstein S.**, 1988, Molecular cloning of complementary DNA encoding maize nitrite reductase. Molecular analysis and nitrate induction. *Plant Physiol*, **88**: 741-746.
- Lancaster J.R., Vega J.M., Kamin H., Orme-Johnson N.R., Orme-Johnson W.H., Krueger R.J., Siegel L.M.**, 1979, Identification of the iron-sulfur center of spinach ferredoxin-nitrite reductase as a tetranuclear center, and preliminary EPR studies of mechanism. *J Biol Chem*. **25**:1268–1272.
- Laskowski R.A., MacArthur M.W., Moss D.S., Thornton J. M.**, 1993, PROCHECK: a program to check the stereochemical quality of protein structures, *J Appl Cryst*, **26**:283-291.
- Lattman E.**, 2004, The state of the Protein Structure Initiative, *Proteins*, **54**: 611-615.
- Laura Lerner., David A., Horita.,** 1993, Teaching High-Resolution Nuclear Magnetic Resonance to Graduate Students in Biophysics, *Biophysical Journal*, **65**:2692-2697.
- Laskowski R.A., Rullmann J.A., MacArthur M.W., Kaptein R., Thornton J.M.**, 1996, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J. Biomol, NMR*, **8**: 477-486.
- Lerner L. ,. Horita D.A .,** 1993, Teaching high-resolution nuclear magnetic resonance to graduate students in biophysics. *Biophysical Journal*. **65**: 2692-2697.

- Lesk A.M., Chothia C.**, 1980, How different amino acid sequences determine similar protein structures: The structure and evolutionary dynamics of the globins, *J Mol Biol*, **136**:225–70.
- Levitt M.**, 1992, Accurate modelling of protein conformation by automatic segment matching, *J Mol Biol*, **226**:507–533.
- Li Z., Jeffrey S.**, 1998, what should the Z-score of native protein structures be, *Protein Science*, **7**:1201-1207.
- Lopez-Ruiz A., Verbelen J.P., Roldan J.M., Diez J.**, 1985, Nitrate reductase of green algae is located in the pyrenoid. *Plant Physiol*, **79**: 1006-1010.
- Luque I., Flores E., Herrero A.**, 1993, Nitrite reductase gene from *Synechococcus* sp PCC7942: homology between cyanobacterial and higher plant nitrite reductases. *Plant Mol Biol*, **21**:1201-1205.
- Luthy R., Bowie J.U., Eisenberg D.**, 1992, Assessment of protein models with threedimensional profiles, *Nature*, **356**:83-85.
- Mathilde C.**, 2005, Méthodes de détection des similarités structurales : caractérisation des motifs conservés dans les familles de structures pour l'annotation des génomes, Thèse de Doctorat, Doctorat de l'Université Pierre et Marie Curie, Paris.
- Martin A.C., Orengo C.A., Hutchinson E.G., Jones S., Karimirantzou M., Laskowski R.A.**, 1998, Protein folds and function, *Structure*, **6**:875-84.
- Marti-Renom M.A., Stuart A.C., Fiser A., Sanchez R., Melo F., Sali A.**, 2000, Comparative protein structure modeling of genes and genomes, *Annu. Rev. Biophys. Biomol. Struct*, **29**:291-325.
- McGarvey, P., Huang H., Barker W.C., Orcutt, B.C., Wu C.H.**, 2000, PIR Web site: New resource for bioinformatics, *Bioinformatics*, **16**:290-291.
- McRee D.E., Richardson D.C., Richardson J.S., Siegel L.M.**, 1986, The heme and Fe4S4 cluster in the crystallographic structure of Escherichia coli sulfite reductase, *J. Biol.Chem.* **261**: 10277-10281.
- Merchán F., Prieto R., Kindle K.L., Llama M.J., Serra J.L., Fernández E.**, 1995, Isolation, sequence and expression in Escherichia coli of the nitrite reductase gene from the filamentous, thermophilic cyanobacterium Phormidium laminosum. *Plant Mol Biol*, **27**:1037-1042.
- Meyer J.**, 2001, Ferredoxins of the third kind. *FEBS Letters*, **509**:1-5
- Mikami B., Ida S.**, 1989, Spinach ferredoxin-nitrite reductase: characterization of catalytic activity and interaction of the enzyme with substrates. *J. Biochem.* **105**: 47-50.

- Morigasaki S., Takata K., Suzuki T., Wada K.,** 1990, Purification and characterization of ferredoxin-NADP⁺ oxidoreductase-like enzyme from radish root tissues. *Plant Physiol*, **93**: 896-901.
- Morris A.L., MacArthur M.W., Hutchinson E.G., Thornton J.M.,** 1992, Stereochemical Quality of Protein Structure Coordinates. *Proteins*, **12**: 345-364.
- Moult J., Fidelis K., Rost B., Hubbard T., Tramontano A.,** 2005, Critical assessment of methods of protein structure prediction (casp)–round 6, *Proteins*, **7**:3–7.
- Mulder G.J.,** 1839, On the composition of some animal substances. *Journal für praktische Chemie*, **16**: 129.
- Nakamura, Y., Masuda, K., Umemiya Y.,** 2001, National Institute of Environment and Health Science, NARO, Dept. of Plant, *Cell and Environment*, 305-8605.
- Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S., Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.,** 2002, Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1 (supplement). *DNA Res. Aug* **31**:135-48.
- Needleman S. B., Wunsch C. D.,** 1970, A general method applicable to the search for similarities in the amino acid sequence of two proteins, *J Mol Biol*, **48**:443–453.
- Novotny M., Kleywegt G.J.,** 2005, A survey of left-handed helices in protein structures, *J Mol Biol*, **347**:231-241.
- Ondrias M.R., Carson S.D., Hirasawa M., Knaff D.B.,** 1985, *Biochim. Biophys. Acta* **830**: 159-163.
- Orea A., Pajuelo P., Pajuelo E., Márquez A.J. and Romero J.M.,** 2001. Characterisation and expression studies of a root cDNA encoding for ferredoxin-nitrite reductase from *Lotus japonicus*. *Physiologia Plantarum*, **113**:193-202.
- Orengo C.A., Pearl F., Bray J., Martin A.C., LoConte L., Todd A., Thornton J.M.,** 1999, The CATH Database Provides Insights on Protein Structure/Function Relationships. *Nucleic Acids Res.* **27**:275-279.
- Pauling L., Corey R.B., Branson H.R.,** 1951, The structure of proteins: two hydrogenbonded helical configurations of the polypeptide chain, *Proc Natl Acad Sci USA* **37**:205-211.
- Pauling L., Corey R.B.,** 1951, The pleated sheet, a new layer configuration of polypeptide chains, *Proc Natl Acad Sci USA*, **37**: 251-256.
- Pearl F., Todd A., Sillitoe I., Dibley M., Redfern O., Lewis T., Bennett C., Marsden R., Grant A., Lee D., Akpor A., Maibaum M., Harrison A., Dallman T., Reeves G., Diboun I., Addou S., Lise S., Johnston C., Sillero A., Thornton J.,**

- Orengo C.**, 2005, The cath domain structure database and related resources gene3d and dhs provide comprehensive domain family information for genome analysis, *Nucleic Acids Res*, **33**:247–51.
- Pearson WR.**, 1995, Comparison of methods for searching protein sequence databases, *Protein Sci* **4**:1145-1160.
- Pearson WR.**, 2000, Flexible sequence similarity searching with the FASTA3 program package, *Methods Mol Biol*, **132**:185-219.
- Pearson WR.**, 1990, Rapid and sensitive comparison with FASTA and FASTP, *Methods Enzymol*, **183**:63–98.
- Peakman T., Crouzet J., Mayaux J.F., Busby S., Mohan S., Harborne N., Wootton J., Nicolson R., Cole J.**, 1990, Nucleotide sequence, organisation and structural analysis of the products of genes in the nirB-cysG region of the Escherichia coli K-12 chromosome, *J. Biochem.* **191**:315-323.
- Perutz M.F., Rossmann M.G., Cullis A.F., Muirhead H., Will G., North A.C.T.**, 1960, Structure of hemoglobin, *Nature*, **185**:416-422.
- Plattner J.J., Greer J., Fung A.K., Stein H., Kleinert H.D., Sham H.L., Smital J.R., Perun T.J.**, 1986, Peptide analogues of angiotensinogen. Effect of peptide chain length on renin inhibition, *Biochem biophys res commun*, 139-982.
- Rafael Ördög.**, 2008, Pydet, a pymol plug-in for visualizing geometric concepts around proteins. *Bioinformatics*, **2**:346–347.
- Ramachandran G.N., Ramakrishnan C., Sasisekharan V.**, 1963, Stereochemistry of polypeptide chain configurations, *Journal of Molecular Biology*, **7**:95–99.
- Gibas C., Jambeck P.**, 2002, Introduction à la bioinformatique. O'Reilly, Paris, 1st edition.
- Ramachandran G.N., Ramakrishnan C., Sesisakharan V.**, 1963, Stereochemistry of polypeptide chain configurations, *J Mol Biol*, **7**: 95–99.
- René heller., Robert esnault., claude lance.**, 1998, physiologie végétale. DUNOD. Paris .6 edn 160p ISBN 2 10 003991 1.
- Rideout J.H., Chaillou S., Raper D.**, 1994, Ammonium and nitrate uptake by soybean during recovery from nitrogen deprivation, *J.Exp.Bot*, **45**: 23-33.
- Reichhardt T.**, 1999, its sink or swim as a tidal wave of data approaches, *nature*, **399**:517-520.
- Roberto S., Andrej S.**, 2000, Comparative Protein Structure Modeling Introduction and Practical Examples with Modeller, *Methods in Molecular Biology*, **143**:97-129.
- Ross-Macdonald P., Sheehan A., Friddle C., Roeder G.S., Snyder M.**, 1999, Transposon mutagenesis for the analysis of protein production, function, and localisation, *Methods enzymol*, **303**:512-32.

Rost B., 1999, Twilight zone of protein sequence alignments, *Protein Eng*, **12**:85–94.
Rychlewski L., Zhang B., Godzik A., 1998, Fold and function predictions for Mycoplasma genitalium proteins, *Folding Design*, **3**:229–38.

Sali A., Blundell T.L., 1993, Comparative protein modelling by satisfaction of spatial restraints, *J Mol Biol*, **234**:779-815.

Sali A., 1995, Modelling mutations and homologous proteins, *Curr Opin Biotech*, **6**:437–51.

Sánchez R., Sali A., 1997, Advances in comparative protein-structure modeling, *Curr Opin Struct Biol*, **7**:206–14.

Saqi M.A.S., Russell R.B., Sternberg M.J.E., 1999, Misleading local sequence alignments: implications for comparative protein modelling, *Protein Eng*, **11**:627–30.

Sander L., Jensen P.E., Back L.F., Stummann B.M., Henningsen K.W., 1995, Structure and expression of a nitrite reductase gene from bean (*Phaseolus vulgaris*) and promoter analysis in transgenic tobacco. *Plant Mol. Biol.* **27**: 165-177.

Sandusky K.V., Fuhrmann M., Nguyen J., Utterback D., Saudek T.R., Phillips D.M., Merrick C.A., Tomb J.M., Dougherty J.F., Bott B.A., Hu K.F., Lucier P.C., Peterson T.S., Smith S.N., Hutchison H.O., C.A., Venter J.C., 1995, The minimal gene complement of Mycoplasma genitalium. *Science*, **270**: 397-403.

Schmitz S., Bohme H., 1995, Amino acid residues involved in functional interaction of vegetative cell Ferredoxin from the *Cyanobacterium anabaena* SP. PCC 7120 with ferredoxin:nad p reductase, nitrite reductase and nitrate reductase . *Biochimica et biophysica acta. Bioenergetics.* **1231**:335-341.

Serra J.L., Ibarlucea J.M., Arizmendi J.M., Llama M.J., 1982, Purification and properties of the assimilatory nitrite reductase from barley *Hordeum vulgare* leaves. *Biochem. J.* **201**: 167-170.

Siegel L.M., Wilkerson J.Q., 1989, Structure and function of spinach ferredoxin-nitrite reductase. In *Molecular and Genetic Aspects of Nitrate Assimilation*, J. L. Wray and J. R. Kinghorn, eds., *Oxford Science.* 263–283.

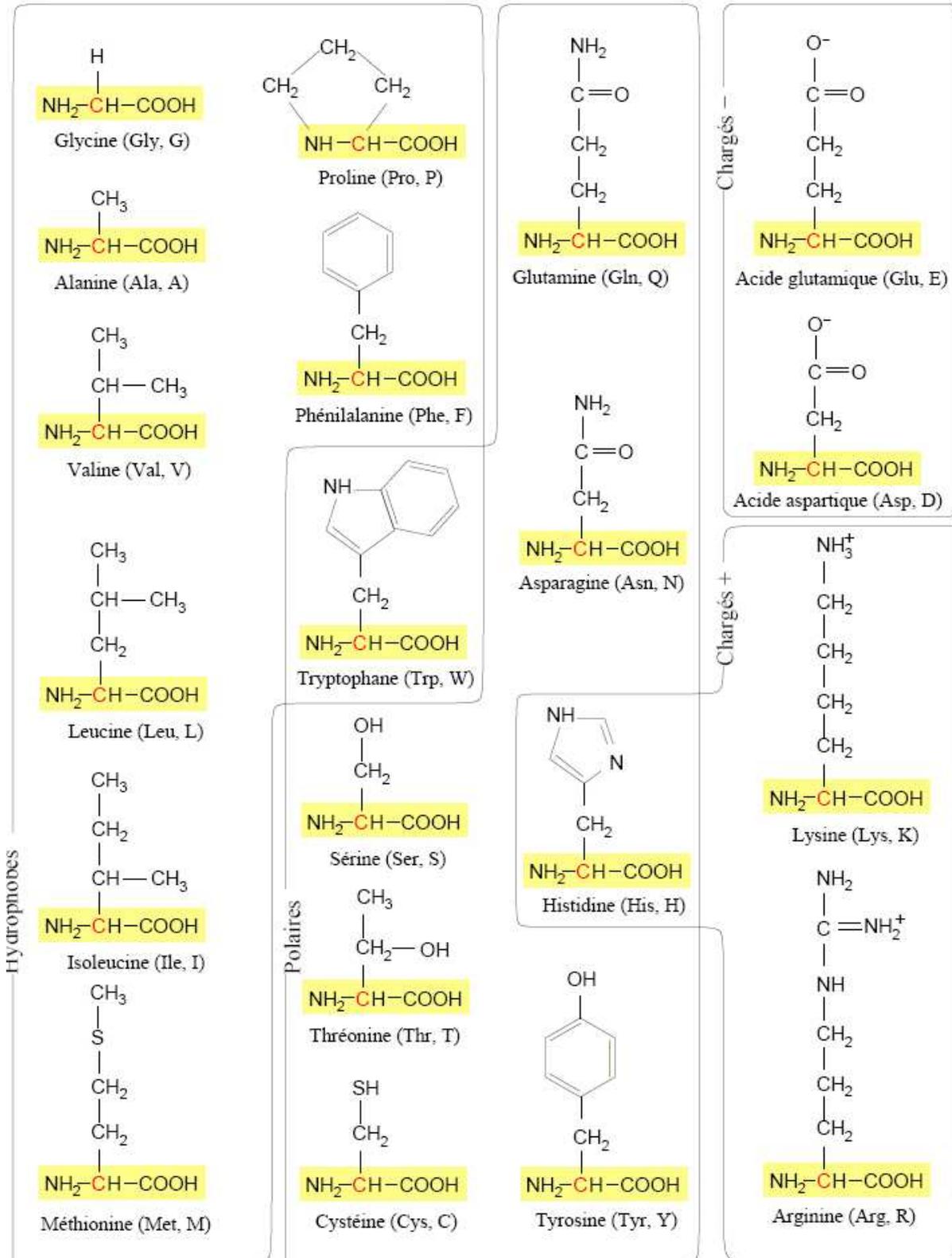
Shirley Hui., 2005, FlexSADRA: Flexible Structural Alignment using a Dimensionality Reduction Approach.. thesis.

Sippl M.J., 1993, Recognition of errors in three-dimensional structures of proteins, *Proteins*, **17**:355-362.

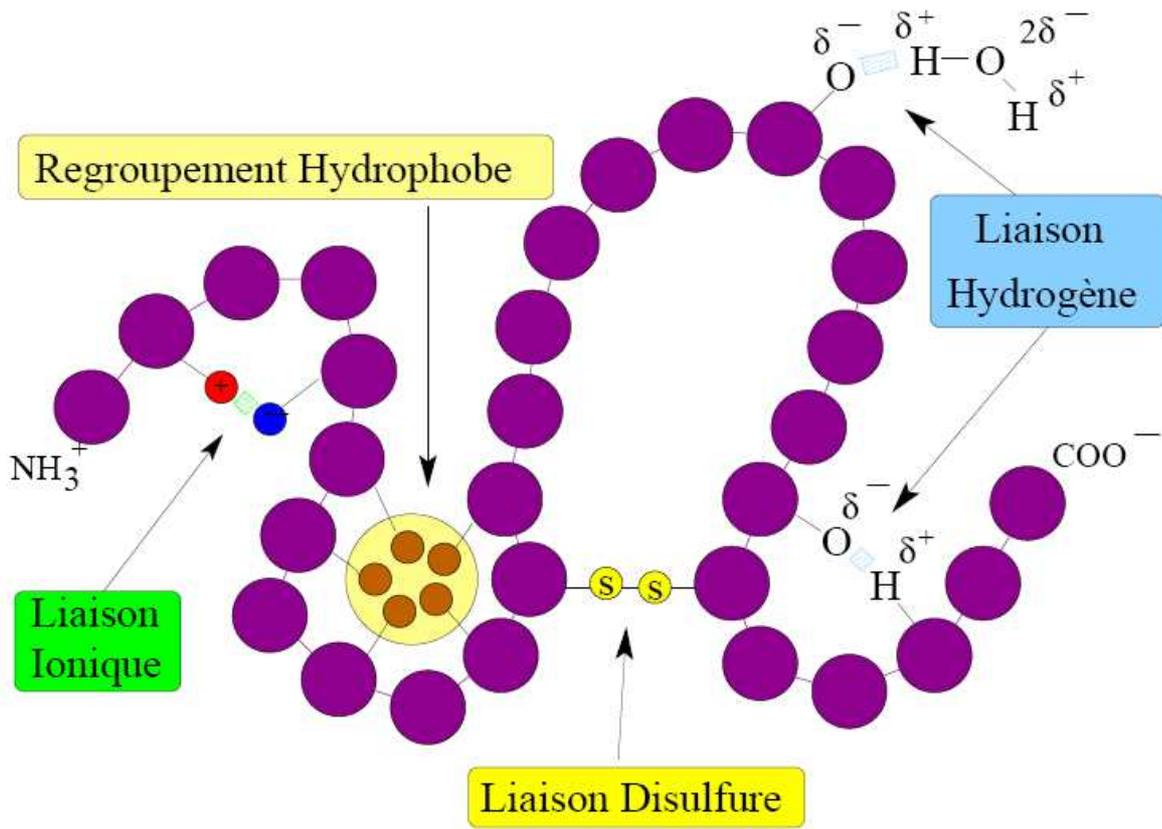
Siegel L.M., Rueger D.C., Barber M.J., Krueger R.J., Orme-Johnson N.R., Orme-Johnson W.H., 1982, scherichia coli sulfite reductase hemoprotein subunit. Prosthetic groups, catalytic parameters, and ligand complexes. *J. Biol. Chem.* **257**: 6343-6350.

- Small I.S., Gray J.C.**, 1984, Synthesis of wheat leaf nitrite reductase de novo following induction with nitrate and light. *Eur. J. Biochem.* **145**: 291-297.
- Smith T.F., Waterman M.S.**, 1981, Identification of common molecular subsequences, *J. Mol. Biol.* **147**:195-197.
- Smyth M.S., Martin J.H.**, 2000, X ray crystallography, *J Clin Pathol Mol Pathol*, **53**:8-14.
- Stryer L.**, 1981, Biochemistry. 2nd ed. W. H. Freeman and Co., San Francisco. (QP514.2.S66).
- Suzuki A., Oaks A., Jacquot J-P., Vidal J., Gadal P.**, 1985, An electron transport system in maize roots for reactions of glutamate synthase and nitrite reductase. Physiological and immunochemical properties of the electron carrier and pyridine nucleotide reductase. *Plant Physiol* **78**:374-378.
- Swamy U., Wang M., Tripathy J.N., Kim S.K., Hirasawa M., Knaff D.B., Allen J.P.**, 2005, Structure of Spinach Nitrite Reductase: Implications for Multi-electron Reactions by the Iron-Sulfur:Siroheme Cofactor, *Biochemistry*, **44**: 16054-16063.
- Tatusov R.L., Mushegian A.R., Bork P., Brown N.P., Hayes W.S., Borodovsky M.**, 1996, Metabolism and evaluation of *Haemophilus Infulenzae* deduced from a whole-genome comparison with Echerichia coli, *Curr biol*, **6**: 279-91.
- Tatusov R.L., Koonin E.V., Lipman D.J.**, 1997, A genomic perspective on protein families, *Science*, **278**:631_7.
- Taylor W.R., Hatrick K.**, 1994, Compensating changes in protein multiple sequence alignments, *Protein Eng*, **7**:341-348.
- Taylor W.R.**, 1996, Multiple protein sequence alignment: algorithms and gap insertion, *Methods Enzymol*, **266**:343-67.
- Terada Y., Aoki H., Tanaka T., Morikawa H., Ida S.**, 1995, Cloning and nucleotide sequence of a leaf ferredoxin-nitrite reductase cDNA of rice. *Biosci. Biotechnol. Biochem.* **59**: 2183-2185.
- Thompson J.D., Higgins D.G., Gibson, T.J.**, 1994, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-speci_c gap penalties and weight matrix choice, *Nucleic Acids Res*, **22**:4673-4680.
- Tramontano., V Morea.**, 2003, Assessment of homology-based predictions in casp5, *Proteins*, **6**:352-68.
- Uetz P., Giot L., Cagney G., Mansfield T.A., Judson R.S., Knight J.R.**, 2000, A comprehensive analysis of protein-protein interactions in *Saccharomyces Cerevisiae*, *Nature* **403**:623-7.

- Uma S., Wang M., Tripathy J.N., Kim S.K., Hirasawa M., Knaff D.B., Allen J.P.,** 2005 Structure of spinach nitrite reductase: implications for multi-electron reactions by the iron-sulfur:siroheme cofactor. *Biochemistry*, **44**:16054-63.
- Vega J.M., Menacho A., León J.,** 1991, Nitrate assimilation by microalgae, *Trends Photochem, Photobiol*, **2**:11–69.
- Vega J.M., Kamin H.,** 1977, Spinach nitrite reductase: Purification and properties of a siroheme-containing iron-sulfur enzyme, *J Biol Chem*, **252**:896-909.
- Vitkup D., Melamud E., Moul J., Sander C.,** 2001, Completeness in structural genomics. *Nat Struct Biol*, **8**:559–66.
- Warren M.J., Roessner C.A., Santander P.J., Scott A.I.,** 1990, The Escherichia coli cysG gene encodes S-adenosylmethionine-dependent uroporphyrinogen III methylase. *Biochem. J.* **265**:725–729.
- Watson J.D., Crick F.H.C.,** 1953, A structure for deoxyribose nucleic acid, *Nature* **171**:737-738.
- Weber I.T., Miller M., Jaskolski M., Leis J., Skalla A.M., Wlodawer A.,** 1989, Molecular Modelling of HIV-1 protease and its substrate binding site, *science*, 243-928.
- Westbrook J., Yang H., Feng Z., Berman H.M.,** 2005, The use of mmCIF architecture for PDB data management. In: Hall S.R., McMahon B., editors. *International Tables for Crystallography*. Vol. G. The Netherlands: Springer, Dordrecht, pp. 539–543.
- White O., Eisen J.A., Heidelberg JF, Hickey EK, Peterson JD, Dodson RJ, Haft DH, Gwinn ML, Nelson WC, Richardson DL, Moffat KS, Qin H, Jiang L, Pamphile W, Crosby M, Shen M, Vamathevan JJ, Lam P, McDonald L, Utterback T, Zalewski C, Makarova KS, Aravind L, Daly MJ, Minton KW, Fleischmann RD, Ketchum KA, Nelson KE, Salzberg S, Smith HO, Venter JC, Fraser CM.,** 1999, Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1. *Science*. **19**:1571–1577.
- Wilkerson J.O., Janick P.A., Siegel L.M.,** 1983, Electron paramagnetic resonance and optical spectroscopic evidence for interaction between siroheme and tetranuclear iron-sulfur center prosthetic groups in spinach ferredoxin-nitrite reductase. *Plant physiology and biochemistry*. **22**:5048-5054.
- Ye Y., Godzik A.,** 2003, Flexible structure alignment by chaining aligned fragment pairs allowing twists, *Bioinformatics*, **19**:246–255.
- Zhang Z.T.,** 1997, Relations of the numbers of protein sequences, families and folds, *Protein Eng*, **10**:757–61.



الشكل (1): البنية الجزيئية للأحماض الأمينية



الشكل (2): يمثل الروابط الكيميائية المشكلة للبنية الثالثية للبروتين