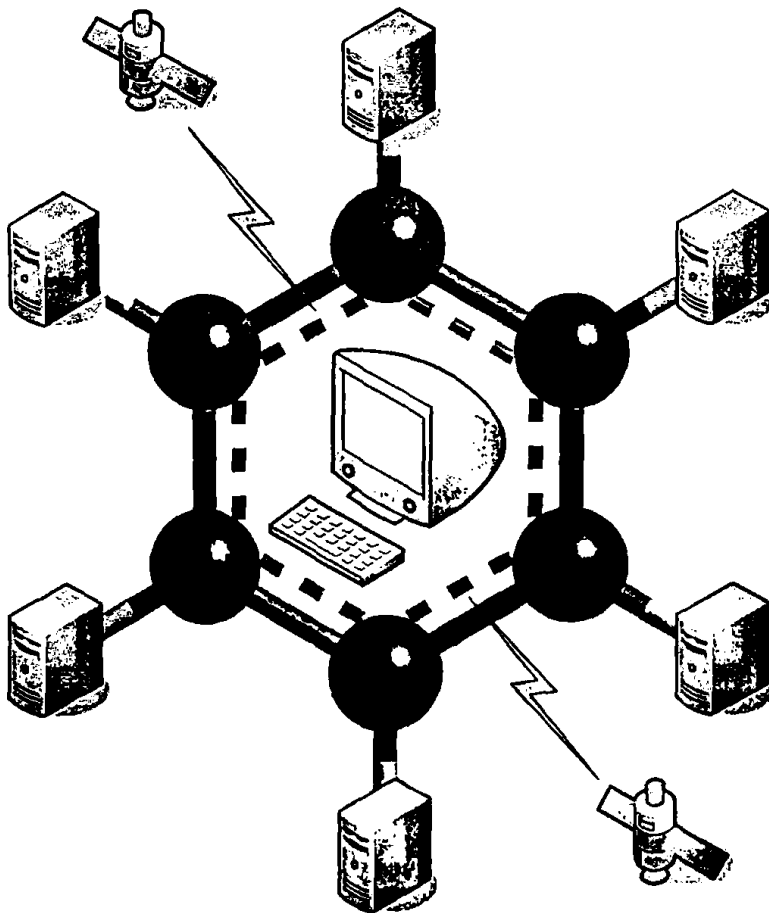


ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΧΗΜΕΙΑΣ

42

Θεοφάνης Γ. Καλατζής
Μαθηματικός

**Συμβολή στην ενοποίηση μεγάλης κλίμακας λογισμικών
προσομοιώσεων (software integration) με εφαρμογές
στην βιοπληροφορική**



ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΙΩΑΝΝΙΝΑ 2007



Στους γονείς
και στα Αδέρφια μου

Ευχαριστίες

Η διατριβή αυτή είναι το αποτέλεσμα μιας επιτυχούς προσπάθειας, για την ολοκλήρωση της οποίας θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή του Χημικού Τμήματος του Πανεπιστημίου Ιωαννίνων κ. Ιωάννη Δημητρόπουλο για τον χρόνο, τις πολύτιμες γνώσεις και την ηθική υποστήριξη που προσέφερε, καθώς και τα δύο μέλη της τριμελούς επιτροπής, καθηγητή κ. Ισαάκ Λαγαρή του Τμήματος Πληροφορικής και καθηγητή κ. Ανδρέα Ράπτη του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων.

Κατά την προσπάθεια αυτή συνέβαλλαν εποικοδομητικά, ο καθηγητής του Τμήματος Μαθηματικών του Πανεπιστημίου Ιωαννίνων κ. Ανδρέας Λεοντίσης τον οποίο ευχαριστώ θερμά, καθώς και ο καθηγητής του Τμήματος Πληροφορικής του Πανεπιστημίου Ιωαννίνων κ. Δημήτριος Φωτιάδης ο οποίος στήριξε συμβουλευτικά, ηθικά και επιστημονικά την προσπάθεια αυτή.

Επίσης, θα ήθελα να ευχαριστήσω τον καθηγητή του Τμήματος Επιστήμης και Τεχνολογίας Υλικών του Πανεπιστημίου Ιωαννίνων κ. Δημήτρη Παπαγεωργίου, ο οποίος όλα αυτά τα χρόνια προσέφερε τις γνώσεις και την εμπειρία του σε όλα τα θέματα που αφορούν τις τεχνολογίες της διατριβής.

Τέλος, θα ήθελα να ευχαριστήσω την Ξένια για την υπομονή και ηθική υποστήριξή της, καθώς και όλους τους φίλους και συνεργάτες μου.



Αρ. 816:.....1426.....200.7,,

at ...
...



Συντομογραφίες

AO	atomic orbital
CAS	complete active space
CASSCF	complete active space self-consistent-field
CCSD	coupled-cluster theory with single and double excitations
CI	configuration interaction
CISD	configuration interaction with single and double excitations
CLSCF	closed-shell self-consistent-field
CSF	configuration state function
DFT	density functional theory
DIIS	Direct Inversion in the Iterative Subspace
GTO	Gaussian-type orbital
GVB	generalized valence bond
HF	Hartree-Fock
HOMO	highest occupied molecular orbital
IP	ionization potential
LCAO	linear combination of atomic orbitals
LUMO	lowest unoccupied molecular orbital
MO	molecular orbital
MCSCF	multiconfiguration self-consistent-field
MP2	second-order Møller-Plesset perturbation theory
NBO	natural bond orbital
PES	potential energy surface
RHF	restricted Hartree-Fock
ROHF	restricted open shell Hartree-Fock
SCF	self-consistent-field
STO	Slater-type orbital
TZ	triple zeta
UHF	unrestricted Hartree-Fock
ZPE	zero point energy



Πρόλογος

Η παρούσα διατριβή, με τίτλο «Συμβολή στην ενοποίηση μεγάλης κλίμακας λογισμικών προσομοιώσεων (software integration) με εφαρμογές στην βιοπληροφορική», αποτελεί μια συμβολή στην περαιτέρω ανάπτυξη της μοριακής κβαντικής προσομοίωσης, η οποία έχει υψηλές απαιτήσεις σε υπολογιστικούς πόρους.

Ο κύριος στόχος είναι η ανάπτυξη κατάλληλων εφαρμογών και μεθόδων ώστε να επιτευχθούν προσομοιώσεις μεγάλης κλίμακας σε συνδυασμό με θεωρίες ικανές να «εξαντλήσουν» την διαθέσιμη υπολογιστική ισχύ και να εξάγουν αξιόπιστα αποτελέσματα. Όλα τα συστήματα που αναπτύχθηκαν υποστηρίζουν παραλληλοποιημένη λειτουργία για τον υπολογισμό των ab-initio ιδιοτήτων. Επιπλέον, για όλα σχεδόν τα αποτελέσματα της παρούσας διατριβής εκτελέστηκαν παράλληλοι υπολογισμοί σε συστοιχίες υπολογιστών (computer clusters) καθώς και σε υπολογιστικές αρχιτεκτονικές με κοινή μνήμη (shared memory machines). Μερικές από τις νέες διαδικασίες και μεθόδους που αναπτύχθηκαν στην διατριβή έχουν ως άμεσο στόχο την εφαρμογή τους σε μεγάλα μοριακά συστήματα βιολογικής σημασίας.

Αναλυτικότερα, στην διατριβή αυτή ολοκληρώθηκαν τα εξής:

✓ Αναπτύχθηκε το ενοποιημένο προγραμματιστικό περιβάλλον MERGAM, το οποίο αφορά την διασύνδεση των πακέτων GAMESS-US και MERLIN/MCL. Το MERGAM προσφέρει, για πρώτη φορά, νέες δυνατότητες για την εξερεύνηση της διαμοριακής υπερεπιφάνειας δυναμικής ενέργειας μέσω αλγορίθμων βελτιστοποίησης και ab-initio κβαντικών υπολογισμών. Το σύστημα παρέχει την δυνατότητα για να αναπτυχθούν αυτοματοποιημένες στρατηγικές εξερεύνησης της υπερεπιφάνειας δυναμικής ενέργειας, μέσω της γλώσσας υψηλού επιπέδου MCL που περιλαμβάνει το πακέτο MERLIN. Στο ενοποιημένο σύστημα, αναπτύχθηκε κώδικας για τον επακριβή υπολογισμό της διαμοριακής ενέργειας αλληλεπίδρασης. Το σύστημα εφαρμόστηκε στην μελέτη της ασθενούς διαμοριακής αλληλεπίδρασης τύπου OH...π στο διμερές βενζόλιο και νερό.

✓ Αναπτύχθηκε το σύστημα TINGAM, το οποίο εκτελεί ab-initio μοριακή δυναμική προσομοίωση, με την διασύνδεση των πακέτων GAMESS-US και TINKER. Το νέο αυτό σύστημα παρέχει υπολογισμό των δυνάμεων (forces) με ab-initio κβαντικές μεθόδους (on-the-fly), ενώ η κίνηση των πυρήνων περιγράφεται από την κλασική μηχανική. Το σύστημα αξιολογήθηκε επιτυχώς με εφαρμογές σε πεπτίδια και μόρια



βιολογικής/βιοτεχνολογικής σημασία όπως οι μιμητές σερινοπρωτεασών. Ο σχεδιασμός των μιμητών αυτών έγινε με το εργαλείο LYSTAS το οποίο αναπτύχθηκε ως επέκταση του TINKER.

✓ Αναπτύχθηκε νέα μεθοδολογία για τον υπολογισμό της διαμοριακής ενέργειας αλληλεπίδρασης σε OH..π ετεροδιμερή μοριακά συστήματα κάνοντας χρήση πολυ-οριζουσιακών κυματοσυναρτήσεων (CASSCF/NEVPT2). Η νέα υπολογιστική μεθοδολογία στοχεύει στην όσο τον δυνατόν ακριβέστερη περιγραφή της διαμοριακής επιφάνειας δυναμικής ενέργειας σε ετεροδιμερή μοριακά συστήματα που αλληλεπιδρούν με ασθενείς δυνάμεις τύπου van der Waals και δεν μπορούν να περιγραφούν επαρκώς από μονο-οριζουσιακές κυματοσυναρτήσεις, όπως η μέθοδος Hartree Fock. Η νέα αυτή υπολογιστική διαδικασία είναι γενική και δεν περιορίζεται μόνο σε OH..π συστήματα, τα οποία χρησιμοποιήθηκαν ως μοντέλα για να δημιουργηθεί η νέα μεθοδολογία.



ΠΕΡΙΕΧΟΜΕΝΑ

Εισαγωγή.....	13
---------------	----

ΚΕΦΑΛΑΙΟ 1 – Θεωρία Κβαντικής Προσομοίωσης

1.1. Θεμελιώδης θεωρία Μοριακής Κβαντικής Προσομοίωσης.....	24
1.1.1. Η προσέγγιση Born-Oppenheimer	25
1.1.2. Η μέθοδος Hartree-Fock και η θεωρία SCF (Self Consistent Field)....	27
1.1.2.1. Εξισώσεις Roothaan - Επίλυση της Hartree Fock	30
1.1.2.2. Αξιολόγηση Της Μεθόδου Hartree- Fock.....	32
1.1.2.3. Σφάλματα της Μεθόδου HF.....	32
1.1.3. Επιλογή Βάσεων.....	33
1.1.3.1. Κατηγορίες συνόλων βάσεων (basis sets)	33
1.1.3.1.1. Τροχιακά τύπου Slater (STO).....	33
1.1.3.1.2. Τροχιακά τύπου Gaussian (GTO).....	35
1.1.3.2. Διαφορές LC-GTO SCFMO και LC-STO SCFMO.....	36
1.1.3.3. Απλοποιήσεις των GTO	36
1.1.3.4. Ελάχιστο Βασικό Σύνολο.....	36
1.1.3.5. Βάση Διπλού Ζήτα (DZ)	37
1.1.3.6. Σύνολα βάσεων Split-Valence	37
1.1.3.6.1. Βάση τύπου STO-NG	38
1.1.3.6.2. Βάση τύπου STO-3G.....	38
1.1.3.6.3. Βάση τύπου 4-31G.....	38
1.1.3.6.4. Βάση τύπου 6-31G*.....	39
1.1.3.7. Slater ή Gaussian	39
1.1.4. Μέθοδοι υπολογισμού της ηλεκτρονιακής συσχέτισης	39
1.1.4.1. Θεωρία των διαταραχών Moller- Plesset	40
1.1.4.2. Μέθοδοι coupled cluster (CC).....	41
1.1.4.2.1. Μέθοδος coupled cluster doubles (CCD).....	42
1.1.4.2.2. Οι μέθοδοι coupled cluster singles, doubles και triples (CCSD και CCSD(T))	42
1.1.4.3 Quadratic Configuration Interaction με απλές και διπλές υποκαταστάσεις (QCISD)	43
1.1.5. Κλιμάκωση μεθόδων.....	44
1.2. Υπολογιστικές μέθοδοι γεωμετρικής βελτιστοποίησης.	45
1.2.1. Τοπικές μέθοδοι ελαχιστοποίησης (local methods).....	45
1.2.2. Ελαχιστοποιήσεις χωρίς τη χρήση παραγώγων	45
1.2.3. Ελαχιστοποιήσεις με τη χρήση παραγώγων	45
1.2.3.1. Μέθοδοι παραγώγων.....	45
1.2.3.2. Τοπικοί ελαχιστοποιητές με χρήση παραγώγων κατηγορίας conjugate gradient (συζύγων διευθύνσεων).....	46
1.2.3.3. Ελαχιστοποίηση με τη χρήση της μεθόδου Newton	47
1.2.3.3.1. Τοπικοί ελαχιστοποιητές κατηγορίας Newton - Raphson	47
1.2.3.3.2. Μέθοδοι quasi - Newton.....	48

ΚΕΦΑΛΑΙΟ 2 – MERGAM: Ένα προγραμματιζόμενο περιβάλλον βελτιστοποίησης με εφαρμογές στην μελέτη της διαμοριακής ενέργειας αλληλεπίδρασης.



2.1. Εισαγωγή	50
2.1.1. Το πακέτο GAMESS-US	52
2.1.2. Το περιβάλλον μη-γραμμικής βελτιστοποίησης MERLIN.....	53
2.2. Διασύνδεση και αναλυτική περιγραφή του συστήματος MERGAM	54
2.2.1. Η αρχιτεκτονική και η επικοινωνία του συστήματος	55
2.2.2. Ο κυρίως κώδικας του συστήματος	58
2.3. Εφαρμογή του περιβάλλοντος MERGAM σε εξειδικευμένα προβλήματα κβαντικής μοριακής μηχανικής.....	61
2.3.1. Η διαμοριακή ενέργεια αλληλεπίδρασης	61
2.3.1.1. Θεωρητικές μέθοδοι υπολογισμού της διαμοριακής ενέργειας αλληλεπίδρασης	63
2.3.1.2 Η μέθοδος SAPT	64
2.3.1.3 Η μέθοδος της υπερμοριακής προσέγγισης.....	65
2.3.1.4. Το πρόβλημα του BSSE.....	66
2.3.2. Η μέθοδος των Boys και Bernardi.....	68
2.3.2.1 Αναλυτικές παράγωγοι των CP-διορθωμένων ενεργειών	70
2.4. Η υλοποίηση της μεθόδου των Boys και Bernardi στο MERGAM.	71
2.5. Εφαρμογές με το σύστημα MERGAM	73
2.5.1. CP-διορθωμένη γεωμετρική βελτιστοποίηση στο διμερές HF-H ₂ O	73
2.5.2. Εξερεύνηση μιας περιοχής της CP-διορθωμένης διαμοριακής υπερεπιφάνειας δυναμικής ενέργειας του διμερούς C ₆ H ₆ -H ₂ O.....	75
2.5.2.1. Υπολογιστικό κόστος της διερεύνησης	79
2.6. Τεχνικές οδηγίες εκτέλεσης υπολογισμών χρησιμοποιώντας το MERGAM.....	79
2.7. Συμπεράσματα.	80

ΚΕΦΑΛΑΙΟ 3 - Υπολογιστική προσομοίωση διαμοριακών ενεργειών αλληλεπίδρασης τύπου OH..n με την μέθοδο CASSCF/NEVPT2.

3.1. Εισαγωγή	84
3.2. Δυναμική και μη-Δυναμική Ηλεκτρονική Συσχέτιση	85
3.3. Η μέθοδος CI (Configuration Interaction).....	87
3.4. Η μέθοδος MCSCF	88
3.4.1. Υπολογισμός της MCSCF ενέργειας	89
3.4.2. Σύγκλιση της MCSCF	90
3.4.3. Απόδοση της MCSCF	90
3.4.4. Περιορίζοντας το μέγεθος της MCSCF	91
3.5. Η μέθοδος CASSCF(n,m)	92
3.5.1. Οι υπολογισμοί με την μέθοδο CASSCF	93
3.5.2. Η εφαρμογή της μεθόδου CASSCF	93
3.5.3. Σχεδιάζοντας τον κατάλληλο ενεργό χώρο	94
3.6. Θεωρία διαταρακών 2ης τάξης στην CASSCF.....	96
3.6.1. CASPT2.....	96
3.6.2. MRMP/MCQDPT	97
3.6.3. Προβλήματα των CASPT2 και MRMP/MCQDPT	97
3.6.4. NEVPT2	98
3.7. Ανάπτυξη CASSCF μεθοδολογίας με το διμερές C ₂ H ₄ -H ₂ O.....	100
3.7.1. Η επιλογή του ενεργού χώρου CAS.....	100
3.7.1.1. Η διαδικασία αναζήτησης των ενεργών μοριακών τροχιακών	101
3.7.1.2. Δημιουργία των σχηματισμών της ανάπτυξης CI.....	102



3.7.1.3. Τα αποτελέσματα του υπολογισμού CI	103
3.7.1.4. Επεκτείνοντας την βάση στην μέθοδο CI	106
3.7.1.5. Ο ενεργός χώρος για το διμερές C ₂ H ₄ -H ₂ O	106
3.7.1.6. Επεκτείνοντας τον ενεργό χώρο	107
3.7.1.7. Size-extensivity στον ενεργό χώρο	108
3.7.1.8. Η υπολογιστική μεθοδολογία	108
3.7.2. Ο ενεργός χώρος CASSCF(2,2)	110
3.7.2.1. Υπολογισμός της ενέργειας CASSCF(2,2)/6-31+G*	110
3.7.2.2. Ο υπολογισμός της CASSCF(2,2)/6-31+G* διαμοριακής ενέργειας αλληλεπίδρασης	111
3.7.3. Ο ενεργός χώρος CASSCF(10,7)	114
3.7.3.1. Υπολογισμός της ενέργειας CASSCF(10,7)/6-31+G*	114
3.7.3.2. Ο υπολογισμός της CASSCF(10,7)/6-31+G* διαμοριακής ενέργειας αλληλεπίδρασης	116
3.8. Απαλοιφή του BSSE σε υπολογισμούς CASSCF/6-31+G* για το C ₂ H ₄ H ₂ O	118
3.8.1 Υπολογισμός του BSSE με την μέθοδο CASSCF(2,2)/6-31+G*	118
3.8.2 Υπολογισμός του BSSE με την μέθοδο CASSCF(10,7)/6-31+G*	120
3.9. Υπολογισμός του BSSE με την μέθοδο NEVPT2	123
3.9.1. Επέκταση της διαδικασίας σε άλλες βάσεις	124
3.9.2. Επέκταση της μεθόδου για την μελέτη του διμερούς C ₆ H ₆ -H ₂ O	125
3.10. Συμπεράσματα	126

ΚΕΦΑΛΑΙΟ 4 – TINKAM: Δημιουργία ενός ενοποιημένου υπολογιστικού συστήματος με εφαρμογές σε μοριακά συστήματα βιολογικής σημασίας.

4.1. Εισαγωγή	130
4.2. Μοριακή Δυναμική	131
4.2.1. Εφαρμογές της Μοριακής Δυναμικής	131
4.2.2. Εξισώσεις Μοριακής Δυναμικής	132
4.2.3. Τα βήματα μιας προσομοίωσης μοριακής δυναμικής	133
4.3. Το πακέτο Μοριακής Δυναμικής TINKER	133
4.3.1. Το πρόγραμμα μοριακής δυναμικής dynamic του TINKER.	134
4.3.2. Αλγόριθμοι ολοκλήρωσης	135
4.3.2.1. Μέθοδος Verlet Leapfrog	135
4.3.2.2. Ολοκληρωτής ταχυτήτων του Verlet	136
4.3.2.3. Επιλογή του βήματος χρόνου.	137
4.3.2.4. Περιοδικές Οριακές Συνθήκες	137
4.3.3. Θερμοκρασία	138
4.3.3.1. Πως υπολογίζεται η θερμοκρασία	139
4.3.4. Στάδια και διάρκεια των προσομοιώσεων της μοριακής δυναμικής. .	140
4.3.4.1. Το στάδιο εξισορρόπησης	140
4.4. Ab-initio μοριακή δυναμική	142
4.4.1. Born-Oppenheimer μοριακή δυναμική	143
4.5. Η διασύνδεση TINKER και GAMESS-US	144
4.5.1. Η αρχιτεκτονική του συστήματος	144
4.5.2. Οι ρουτίνες που αναπτύχθηκαν στο πακέτο TINKER	147
4.5.3. Η εφαρμογή dynamicab	147
4.5.4. Οι ρουτίνες BEEMAN2 και VERLET2	148
4.6. Εφαρμογές	150



ΠΕΡΙΕΧΟΜΕΝΑ

4.6.1. Το διπεπτίδιο αλανίνης.....	151
4.6.1.1. Υπολογιστικές μέθοδοι προσομοίωσης.....	151
4.6.1.2. Αποτελέσματα από την προσομοίωση AM1 και SM5.42	152
4.6.2. Η κυκλό-εξαγλυκίνη	154
4.6.3. Εκτέλεση προσομοιώσεων με το σύστημα GAMESS-US και TINKER.....	156
4.6.4. Προσομοίωση σε μόρια βιολογικής σημασίας	156
4.7. Συμπεράσματα	157

Συμπεράσματα της διατριβής	159
----------------------------------	-----

ΠΑΡΑΡΤΗΜΑ

Παράρτημα Α.....	161
Παράρτημα Β.....	163
Παράρτημα Γ.....	165
Παράρτημα Δ.....	166
Παράρτημα Ε.....	169
Παράρτημα ΣΤ.....	170
Παράρτημα Ζ.....	178
Παράρτημα Η.....	179
Παράρτημα Θ.....	181
Παράρτημα Ι.....	183
Παράρτημα Κ.....	185
Παράρτημα Μ.....	186
Παράρτημα Ν.....	187

Βιβλιογραφία	189
--------------------	-----



ΕΙΣΑΓΩΓΗ

Η υπολογιστική προσομοίωση (computer simulation) στηρίζεται στην μαθηματική μοντελοποίηση. Η μαθηματική μοντελοποίηση είναι η εφαρμογή κανόνων και μεθόδων για να αναλυθούν πολύπλοκα προβλήματα του φυσικού κόσμου. Ο στόχος είναι να ληφθούν αποφάσεις και να μελετηθούν οι διάφορες πιθανές καταστάσεις ενός φυσικού συστήματος.

Η ιστορία της υπολογιστικής προσομοίωσης ξεκινά κατά την διάρκεια του 2^{ου} παγκοσμίου πολέμου όπου ο μαθηματικός Jon Von Neumann και ο φυσικός Stanislaw Ulam αντιμετώπιζαν το πρόβλημα της συμπεριφοράς των νετρονίων. Ο J V Neumann είχε εμπλακεί με την θεωρία και τον σχεδιασμό πυρηνικών δοκιμών στο εργαστήριο Los Alamos και ήταν ειδικός στην κβαντική θεωρία και την υδροδυναμική, καθώς είχε σπουδάσει χημεία και μαθηματικά, ενώ είχε την ικανότητα να «μεταφράζει» τα προβλήματα των φυσικών σε μαθηματικά μοντέλα, τα οποία μπορούσαν να αναλυθούν και να υπολογιστούν. Το 1947 η ομάδα των Neumann και Ulam με επικεφαλής τον θεωρητικό φυσικό Nicholas Metropolis υλοποίησαν την πρώτη Monte Carlo υπολογιστική προσομοίωση που αφορούσε αλυσιδωτές αντιδράσεις στο εργαστήριο Los Alamos χρησιμοποιώντας έναν υπολογιστή ENIAC¹. Στην δεκαετία του 1950 η υπολογιστική προσομοίωση δεν ήταν τόσο επιτυχής λόγω του μεγάλου υπολογιστικού και οικονομικού κόστους που χρειαζόταν για να εξαχθούν τα αποτελέσματα². Τον Δεκέμβριο του 1961 ο Geoffrey Gordon σχεδίασε ένα εργαλείο για την δημιουργία συστημάτων προσομοίωσης, το οποίο παρείχε πληροφορίες στην εναέρια κυκλοφορία για τον καιρό σε συνεργασία με την IBM η οποία παρείχε το υλικό και το λογισμικό. Οι πρώτες ομάδες προσομοίωσης δημιουργήθηκαν στις εταιρείες: Boeing, Martin Marietta, Air Force Logistics Command, General Dynamics, Hughes Aircraft, Raytheon, Celanese, Exxon, Southern Railway, ενώ οι κατασκευαστές υπολογιστών ήταν οι: IBM, Control Data, National Cash Register, και UNIVAC².

Η υπολογιστική προσομοίωση εφαρμόζεται στις περιπτώσεις όπου:

- Δεν είναι δυνατόν να γίνουν τα πραγματικά πειράματα, όπως π.χ. η περίπτωση του Greenhouse Effect και η γέννηση γαλαξιών.
- Το κόστος σε χρόνο και χρήματα καθώς και ο βαθμός επικινδυνότητας είναι απαγορευτικά όπως η περίπτωση πυρηνικών και πολύπλοκων χημικών αντιδράσεων.



- Το φυσικό σύστημα δεν έχει ακόμα αναπτυχθεί και βρίσκεται στο στάδιο του σχεδιασμού.
- Ζητείται ο έλεγχος διάφορων εναλλακτικών συμπεριφορών της εξέλιξης ενός συστήματος, όπως η πρόβλεψη των καιρικών συνθηκών.

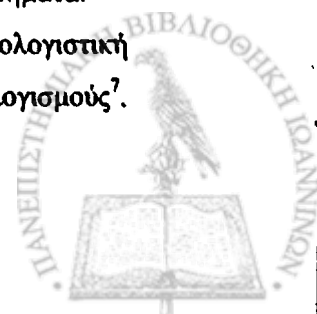
Τα σημαντικά κριτήρια που πρέπει να ληφθούν υπ' όψη πριν οποιασδήποτε υπολογιστική προσομοίωση είναι:

- Το υπολογιστικό κόστος που χρειάζεται να αναπτυχθεί η προσομοίωση.
- Η δυσκολία επαλήθευσης των αποτελεσμάτων λόγω της μη-χρήσης πραγματικών δεδομένων.
- Η προσομοίωση πρέπει να παρέχει όσο το δυνατόν καλύτερες λύσεις αν δεν είναι δυνατόν να παρέχει την βέλτιστη λύση.

Μία από τις πρώτες υπολογιστικές προσομοιώσεις που έγιναν στην ιστορία ήταν η μοριακή προσομοίωση τύπου Monte Carlo, η οποία επιτεύχθηκε στο εργαστήριο Los Alamos το 1953 σε έναν υπολογιστή MANIAC από την ομάδα του Metropolis³. Αμέσως μετά υλοποιήθηκαν υπολογιστικές μοριακές προσομοιώσεις σε μοντέλα που περιγράφονταν: α) από σφαίρες⁴ και β) από σωματίδια που περιγράφονταν από το δυναμικό Lennard-Jones⁵. Αυτές οι πρώτες υπολογιστικές προσομοιώσεις στηρίζονταν στην μοντελοποίηση της αλληλεπίδρασης των σωματιδίων που απαρτίζουν ένα μοριακό σύστημα. Στην συνέχεια άρχισαν να γίνονται λεπτομερείς προσομοιώσεις σε μικρά μοριακά συστήματα και καθώς η υπολογιστική ισχύς εξελισσόταν οι προσομοιώσεις έβρισκαν εφαρμογή σε μεγαλύτερα μοριακά συστήματα όπως πολυμερή και μόρια βιολογικής σημασίας.

Όλες οι αρχικές υπολογιστικές μοριακές προσομοιώσεις αφορούσαν την μελέτη του δυναμικού αλληλεπίδρασης των ατόμων, το οποίο περιγράφεται από την επιφάνεια δυναμικής ενέργειας (Potential Energy Surface – PES). Το 1973 η ομάδα του Clementi ανέπτυξε ένα δυναμικό αλληλεπίδρασης χρησιμοποιώντας ab-initio (Hartree-Fock) ενεργειακά αποτελέσματα από 190 διμερή μόρια του νερού⁶. Στην εργασία αυτή συνδυάστηκαν για πρώτη φορά η κβαντική μηχανική και η στατιστική θερμοδυναμική για την πρόβλεψη μοριακών μακροσκοπικών ιδιοτήτων από καθαρή θεωρία. Η ανάπτυξη και η διαθεσιμότητα των υπολογιστικών πόρων ανέδειξε ακόμα πιο πρακτικούς τρόπους για τον υπολογισμό των επιφανειών δυναμικής ενέργειας, παρόλο που ακόμα και σήμερα εφαρμόζονται μόνο για μικρά μοριακά συστήματα.

Η πρόβλεψη νέων μοριακών ιδιοτήτων μπορεί να γίνει με υπολογιστική προσομοίωση η οποία στηρίζεται σε ακριβείς ab-initio κβαντικούς υπολογισμούς⁷.



Ένας πολύ επιτυχής και ακριβής *ab-initio* υπολογισμός των ενεργειών αλληλεπίδρασης έγινε με υπολογιστική προσομοίωση το 1999⁸.

Η μοριακή προσομοίωση αποτελείται από ένα σύνολο υπολογιστικών τεχνικών και εργαλείων με σκοπό την εξαγωγή και αναπαράσταση των ιδιοτήτων που αφορούν την δομή και τις αντιδράσεις μοριακών συστημάτων χρησιμοποιώντας την τρισδιάστατη μοριακή γεωμετρία. Η μοριακή προσομοίωση έχει πλέον εδραιωθεί ως κλάδος της θεωρητικής χημείας ενώ το 1998 ο John Pople βραβεύτηκε⁹ με το Nobel Χημείας σύμφωνα με την ακόλουθη εξαγγελία των βραβείων Nobel:

John Pople is rewarded for developing computational methods making possible the theoretical study of molecules, their properties and how they act together in chemical reactions. These methods are based on the fundamental laws of quantum mechanics as defined by, among others, the physicist E. Schrödinger. A computer is fed with particulars of a molecule or a chemical reaction and the output is a description of the properties of that molecule or how a chemical reaction may take place. The result is often used to illustrate or explain the results of different kinds of experiment. Pople made his computational technics easily accessible to researchers by designing the GAUSSIAN computer program. The first version was published in 1970. The program has since been developed and is now used by thousands of chemists in universities and commercial companies the world over.

Οι σύγχρονες εφαρμογές της μοριακής προσομοίωσης αφορούν πλέον μοριακά συστήματα βιολογικής σημασίας. Μεγάλη συμβολή στις νέες εφαρμογές έχει η παραλληλοποίηση κώδικα μοριακής προσομοίωσης. Η ανάπτυξη νέων παράλληλων ή κατανεμημένων αλγορίθμων και υπολογιστικών αρχιτεκτονικών παρέχουν υψηλής ακρίβειας προσομοιώσεις σε αρκετά μεγάλα μοριακά συστήματα¹⁰. Στην πλατφόρμα Thunder¹¹ έγινε προσομοίωση 512 μορίων νερού, στην οποία επιτεύχθηκε 83% αξιοπιστία στην παραλληλοποιημένη εκτέλεση μεταξύ 1120 και 3920 επεξεργαστών. Μία άλλη εξαιρετικής σπουδαιότητας υπολογιστική αρχιτεκτονική που χρησιμοποιείται σε μοριακές προσομοιώσεις είναι το σύστημα BlueGene/L¹² της IBM. Το σύστημα αυτό αποτελείται από 65536 κόμβους και 131072 επεξεργαστές ενώ έχει θεωρητική απόδοση 360TFlops/s.

Δεκάδες πακέτα λογισμικού μοριακής προσομοίωσης είναι διαθέσιμα στην επιστημονική και ερευνητική κοινότητα ενώ τα περισσότερα από αυτά διανέμονται



δωρεάν στην ακαδημαϊκή κοινότητα¹³. Τα περισσότερα από αυτά τα πακέτα ξεκίνησαν να αναπτύσσονται την δεκαετία του 1970 ενώ νέες εκδόσεις και βελτιώσεις παρέχονται μέχρι σήμερα. Τα περισσότερα από αυτά τα πακέτα είναι παραλληλοποιημένα ώστε να ανταποκρίνονται σε μεγάλης κλίμακας προσομοιώσεις.

Ανάπτυξη Εφαρμογών Υπολογιστικής Προσομοίωσης

Οι εφαρμογές υπολογιστικής προσομοίωσης απαιτούν την ενοποίηση εργαλείων και τεχνικών που έχουν αναπτυχθεί από πολλαπλούς επιστημονικούς κλάδους¹⁴. Η αναγκαιότητα αυτή δημιουργείται από την αλληλεπίδραση των επιμέρους φυσικών μοντέλων που χρειάζονται για να προσομοιωθεί ένα φυσικό σύστημα, όπως οι κλιματικές αλλαγές. Για τον σκοπό αυτό χρειάζεται να αναπτυχθούν ιδιαίτερες αρχιτεκτονικές λογισμικού οι οποίες θα παρέχουν αξιόπιστη ενοποίηση διάφορων ανεξάρτητων εφαρμογών οι οποίες θα επικοινωνούν μεταξύ τους αποδοτικά, κρύβοντας τεχνικές λεπτομέρειες από τον χρήστη. Η ανάπτυξη αποδοτικού κώδικα για πολύπλοκες προσομοιώσεις μεγάλης κλίμακας σε παράλληλες υπολογιστικές αρχιτεκτονικές είναι γενικά ένα δύσκολο πρόβλημα¹⁵.

Ένα υπολογιστικό σύστημα που αποτελείται από διάφορες ενοποιημένες ανεξάρτητες επιστημονικές εφαρμογές έχει τεράστια εφαρμογή σε όλες τις μεγάλης κλίμακας προσομοιώσεις όπως αυτή που γίνεται στο κέντρο προσομοίωσης μηχανών πυραύλων (CSAR) στο Πανεπιστήμιο του Illinois^{14,16}. Ο στόχος είναι η ανάπτυξη ενός ενοποιημένου υπολογιστικού συστήματος για την πλήρη προσομοίωση των μηχανών των πυραύλων. Το σύστημα αυτό που λέγεται Rocstar έχει επιπλέον εφαρμογές στην προσομοίωση αεροσυμπιεστών, πτερυγίων αεροσκαφών και στην προσομοίωση της ροής του αίματος. Το ενοποιημένο σύστημα προσομοίωσης αποτελείται από τέσσερις ανεξάρτητες κατηγορίες εφαρμογών που είναι:

1. Οι εφαρμογές επίλυσης προβλημάτων φυσικής που είναι γραμμένες σε γλώσσα προγραμματισμού Fortran 90.
2. Οι εφαρμογές διεπικοινωνίας και διαχείρισης εισόδου/εξόδου δεδομένων γραμμένες στην γλώσσα C++.
3. Το περιβάλλον διαχείρισης ολόκληρου του συστήματος προσομοίωσης και
4. Οι μονάδες ελέγχου των εφαρμογών φυσικής και διαχείρισης δεδομένων.

Άλλο ένα σύστημα που αναπτύχθηκε το 1998 αφορά την προσομοίωση της ροής του νερού με την ενοποίηση των πακέτων FENFLOOS και COVISE¹⁷. Πάρα πολλές εφαρμογές υπολογιστικής προσομοίωσης των ρευστών έχουν ενοποιηθεί με το



λογισμικό TENT¹⁸ (TEstbed for Numerical Turbine simulation) το οποίο είναι ανεπτυγμένο σε Java. Επίσης το σύστημα αριθμητικής προσομοίωσης προωθήσεων (Numerical Propulsion System Simulation – NPSS) αναπτύσσεται στην NASA και είναι ένας μεγάλης – κλίμακας υπολογιστικό σύστημα που δημιουργήθηκε από την ενοποίηση ανεξάρτητων εφαρμογών¹⁹.

Η ενοποίηση ανεξάρτητων λογισμικών, με σκοπό την ανάπτυξη συστημάτων προσομοίωσης μεγάλης-κλίμακας, έχει εφαρμογή σε όλο το πεδίο των επιστημονικών και ερευνητικών προσομοιώσεων. Αυτό συμβαίνει γιατί χρήσιμος κώδικας που αναπτύχθηκε ή αναπτύσσεται ακόμα και σήμερα από διάφορα εξειδικευμένα εργαστήρια βρίσκει εφαρμογή σε νέους επιστημονικούς τομείς και νέες ομάδες χρειάζεται να τον ενσωματώσουν σε νέο-δημιουργούμενες μεγάλης κλίμακας εφαρμογές. Στις αρχές της δεκαετίας του 1990 αναπτύχθηκε στο CERN ένα περιβάλλον με σκοπό την ενοποίηση εφαρμογών που είναι γραμμένες στις γλώσσες προγραμματισμού C και Fortran-77²⁰, δεδομένου ότι η βιβλιοθήκη του CERN (CERNLIB) περιέχει εκατοντάδες επιστημονικές εφαρμογές το 80% των οποίων είναι γραμμένες στην γλώσσα προγραμματισμού Fortran. Το περιβάλλον αυτό χρησιμοποιήθηκε για την δημιουργία του πακέτου PAW (Physics Analysis Workstation) το οποίο άρχισε να αναπτύσσεται από το 1988²¹.

Η ανάπτυξη λογισμικού με συνδυασμό των γλωσσών προγραμματισμού C και Fortran έχει μεγάλη εφαρμογή σε πολλά πακέτα επιστημονικών υπολογισμών και προσομοιώσεων. Ειδικότερα, αφορά την παραλληλοποίηση και κατανεμημένη επεξεργασία πολλών εφαρμογών μοριακής προσομοίωσης οι οποίες αναπτύχθηκαν στην γλώσσα προγραμματισμού Fortran 77 και η παραλληλοποίησή τους απαιτεί την χρήση βιβλιοθηκών που είναι γραμμένες στις γλώσσες προγραμματισμού C και C++²². Ένα χαρακτηριστικό παράδειγμα μια τέτοιας διασύνδεσης αφορά το πακέτο κβαντικών υπολογισμών GAMESS-US²³ με το περιβάλλον DDI²⁴ (Distributed Data Interface). Άλλη μία αντίστοιχη προσέγγιση έγινε για τα πακέτα κβαντικών υπολογισμών GAUSSIAN²⁵ και MPQC²⁶. Το πακέτο GAUSSIAN για την παράλληλη επεξεργασία χρησιμοποιεί το περιβάλλον LINDA ενώ το MPQC χρησιμοποιεί την παραλληλοποιημένη βιβλιοθήκη Scientific Computing Toolkit²⁷ που είναι γραμμένη στην γλώσσα C++.

Το πιο κοινό περιβάλλον που διασυνδέθηκε με εφαρμογές κβαντικής μοριακής προσομοίωσης για την παραλληλοποίηση τους είναι το TCGMSG²⁸ (Theoretical Chemistry Group Message-passing toolkit). Αποτελούσε το κύριο σύστημα



παραλληλοποίησης του GAMESS-US μέχρι το 1999 όπου και αντικαταστάθηκε από την βιβλιοθήκη DDI. Στον πίνακα 1 φαίνονται οι βασικές ρουτίνες του TCGMSG σε Fortran και C.

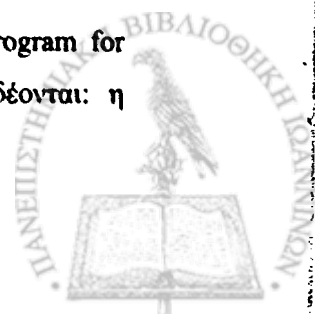
Πίνακας 1

Operation	FORTTRAN	C
Initialization	call PBEGINF()	PBEGIN_(...)
Termination	call PEND()	PEND_()
Identification		
Number of nodes	NNODES()	NNODES_()
Node ID	NODEID()	NODEID_()
Communication Operations		
Send	call SND(...)	SND_(...)
Receive	call RCV(...)	RCV_(...)
Broadcast	call BRDCST(...)	BRDCST_(...)
Synchronization	call SYNCH(...)	SYNCH_(...)
Shared counter	NXTVAL(...)	NXTVAL_(...)
Global operation	call DGOP(...)	DGOP_(...)
Utilities		
Print statistics	call STATS()	STATS_()
Wall time	TCGTIME()	TCGTIME_()

Το πιο σημαντικό χαρακτηριστικό που παρέχει η τεχνολογία διασύνδεσης και ενοποίησης επιστημονικών εφαρμογών είναι η δυνατότητα για ανανέωση ή αναπροσαρμογή μίας επιμέρους εφαρμογής, όταν αυτό κριθεί απαραίτητο, όπως στην περίπτωση της αλλαγής του TCGMSG με το DDI στο πακέτο GAMESS-US.

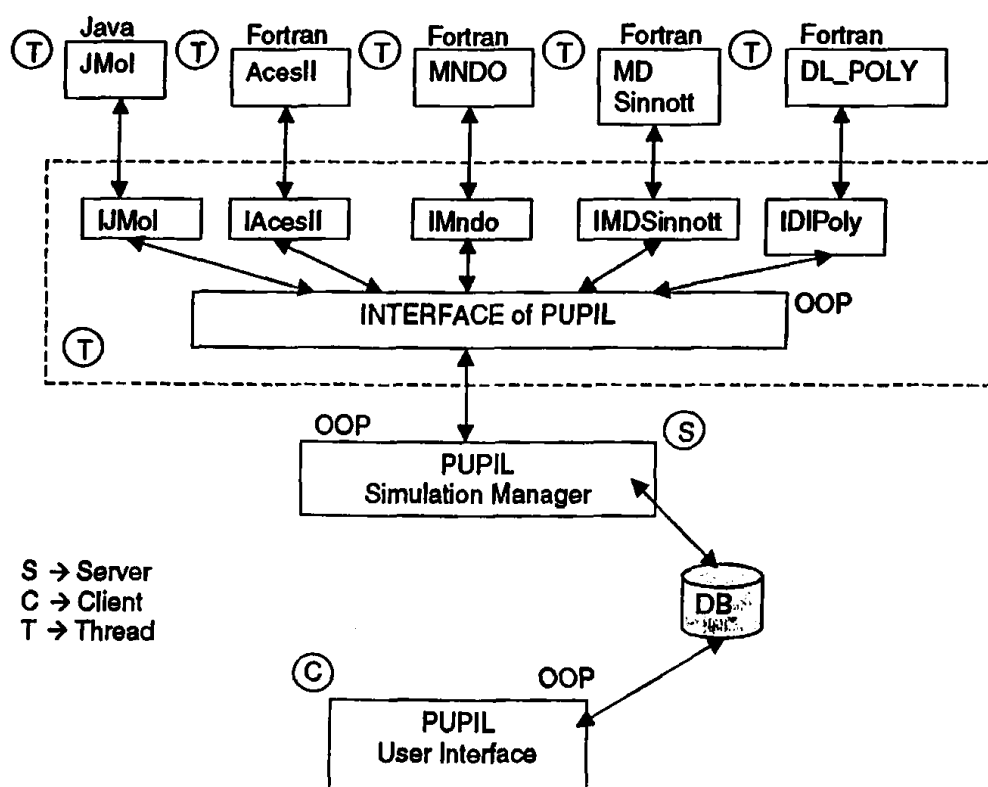
Σε όλα τα πακέτα κβαντικών υπολογισμών και προσομοιώσεων είναι πλήρως διαδεδομένη η διασύνδεση ανεξάρτητων εφαρμογών καθώς και ενοποίηση ανεξάρτητων πακέτων σε ένα κοινό περιβάλλον διαχείρισης τόσο για την παραλληλοποίηση όσο και για την επέκτασή τους με νέες μεθόδους. Ειδικότερα το πακέτο GAMESS-US παρέχει ενσωματωμένες δυνατότητες για την εξωτερική διασύνδεση του με εφαρμογές μοριακής μηχανικής όπως το TINKER²⁹ ή με εξειδικευμένες εφαρμογές κβαντικής προσομοίωσης των ατομικών πυρήνων όπως το υποσύστημα NEO (Nuclear Electronic Orbital)³⁰. Άλλη μία ενοποίηση λογισμικών προσομοίωσης έγινε μεταξύ των πακέτων MOPAC³¹ και VENUS³² και είχε σκοπό την ανάπτυξη ενός συστήματος ab-initio μοριακής δυναμικής³³.

Ένα εξελεγμένο σύστημα μοριακής προσομοίωσης είναι το PUPIL³⁴ (Program for User Packages Interfacing and Linking). Στο σύστημα αυτό διασυνδέονται: η



εφαρμογή μοριακής δυναμικής DL_POLY³⁵ και η εφαρμογή ημι-εμπειρικών κβαντικών υπολογισμών MNDO³⁶. Η διεπικοινωνία μεταξύ των εφαρμογών γίνεται με την τεχνολογία CORBA³⁷, όπου ανταλλάσσονται δεδομένα σχετικά με την διεύθυνση του μοριακού συστήματος, την ενέργεια και τις δυνάμεις σε κάθε βήμα της μοριακής δυναμικής. Το ενοποιημένο σύστημα χρησιμοποιεί την τεχνολογία των wrapper για την αποδοτική διεπικοινωνία και ανταλλαγή δεδομένων μεταξύ των επιμέρους εφαρμογών. Τα δύο κύρια υποσυστήματα του ενοποιημένου περιβάλλοντος μοριακής προσομοίωσης αφορούν διασύνδεση με Java και C καθώς και Fortran με C. Η πλήρης αρχιτεκτονική του συστήματος PUPIL φαίνεται στην εικόνα 1.

Εικόνα 1



Το σύστημα ελέγχθηκε σε εφαρμογές νανοτεχνολογίας (nanorod) όπου προσομοιώθηκε ένα μοριακό σύστημα 108 ατόμων³⁸.

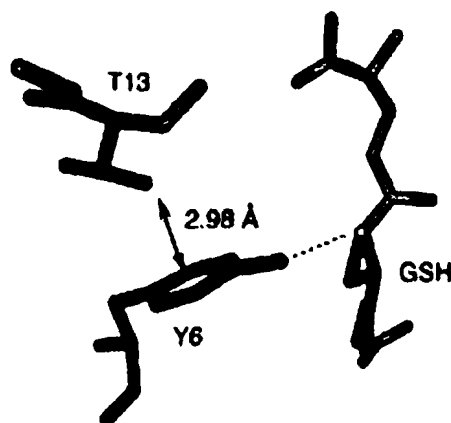
Εφαρμογές και δομή της διατριβής

Οι πιο κοινές κατηγορίες των αποτελεσμάτων που προκύπτουν από την εφαρμογή των πακέτων κβαντικής μοριακής προσομοίωσης μεγάλης κλίμακας αφορούν³⁹:

- Την εύρεση της σταθερή δομής και ενέργειας του μοριακού συστήματος.
- Τον υπολογισμό της ενέργειας δεσμών και αντιδράσεων.
- Την εύρεση της δομής και ενέργειας των μεταπτωτικών καταστάσεων.
- Την μελέτη διαφόρων μηχανισμών αντιδράσεων.
- Τον υπολογισμό των δυναμικών ιονισμού.
- Την κατανομή του φορτίου σε ένα μοριακό σύστημα.
- Τον υπολογισμό των συχνοτήτων δόνησης (IR/RAMAN).
- Τον υπολογισμό των ηλεκτρονικών μεταπτώσεων.

Οι εφαρμογές στις οποίες επικεντρώνονται τα μοντέλα και οι προσομοιώσεις της διατριβής αφορούν ασθενείς διαμοριακές αλληλεπιδράσεις τύπου OH...π, όπως π.χ. εμφανίζονται στο διμερές βενζόλιο-νερό. Οι ασθενείς αλληλεπιδράσεις τύπου OH...π παρατηρούνται κυρίως σε συστήματα που αποτελούνται από μόρια νερού και πρωτεΐνες⁴⁰. Επίσης μόρια νερού που αλληλεπιδρούν με αρωματικούς δακτυλίους των Phe, Trp και Tyr έχουν σημαντική συνεισφορά στην σταθεροποίηση της πρωτεϊνικής δομής⁴¹. Ένα παράδειγμα αλληλεπίδρασης OH...π φαίνεται στην εικόνα 1.2, όπου φαίνεται ένα σύμπλοκο του ένζυμου GSH⁴².

Εικόνα 1.2



OH...π αλληλεπίδραση στο σύμπλοκο glutathione transferase και glutathione (PDB code 6GST).

Η δομή της διατριβής αποτελείται από μία εισαγωγή που αφορά την υπολογιστική προσομοίωση και τέσσερα κεφάλαια ως εξής:

- Στην εισαγωγή γίνεται αναφορά στην επιστημονική υπολογιστική προσομοίωση καθώς και στις μεθόδους ανάπτυξης και ενοποίησης λογισμικού προσομοίωσης.
- Στο πρώτο κεφάλαιο γίνεται αναφορά στην θεωρία των *ab-initio* κβαντικών υπολογισμών καθώς και στις μεθόδους ενεργειακής ελαχιστοποίησης μοριακών συστημάτων που στοχεύουν στην εύρεση της πιο «σταθερής» μοριακής δομής.
- Στο δεύτερο κεφάλαιο παρουσιάζεται η ανάπτυξη ενός ενοποιημένου περιβάλλοντος μοριακής προσομοίωσης διασυνδέοντας για πρώτη φορά τα πακέτα GAMESS-US και MERLIN/MCL. Το σύστημα αναπτύχθηκε με στόχο την επίλυση προβλημάτων κβαντικής χημείας τα οποία απαιτούν ευφυής τεχνικές γεωμετρικής βελτιστοποίησης σε εξειδικευμένες περιοχές της υπερεπιφάνειας δυναμικής ενέργειας μοριακών συστημάτων. Το ενοποιημένο προγραμματιστικό περιβάλλον εφαρμόζεται στην μελέτη των ασθενών διαμοριακών ενεργειών αλληλεπίδρασης σε μικρά μοριακά συστήματα ώστε να εξαχθούν δεδομένα χρήσιμα για την μελέτη μεγαλύτερων συστημάτων.
- Το τρίτο κεφάλαιο αφορά την ανάπτυξη μια νέας μεθοδολογίας για τον υπολογισμό ασθενών διαμοριακών ενεργειών αλληλεπίδρασης με CASSCF/NEVPT2 *ab-initio* μεθόδους που αποτελούν υψηλής ακρίβειας ποιοτική και ποσοτική μοντελοποίηση μοριακών συστημάτων. Οι υπολογιστικές μέθοδοι που εφαρμόζονται έχουν ως κύριο χαρακτηριστικό την πλήρη εκμετάλλευση της διαθέσιμης υπολογιστικής ισχύς, ενώ περιορίζονται σε μικρά μοριακά συστήματα λόγω του πολύ υψηλού υπολογιστικού κόστους.
- Στο τέταρτο κεφάλαιο δίνεται η ανάπτυξη ενός ενοποιημένου περιβάλλοντος με τα πακέτα GAMESS-US και TINKER, το οποίο στοχεύει στην μοριακή δυναμική προσομοίωση με *ab-initio* μεθόδους. Για τον σκοπό αυτό δημιουργείται μία βιβλιοθήκη διασύνδεσης των δύο πακέτων, ώστε να είναι δυνατός ο υπολογισμός της τροχιάς (trajectory) ενός μοριακού συστήματος του οποίου η δυναμική ενέργεια υπολογίζεται με ημι-εμπειρικές και *ab-initio* μεθόδους. Το ενοποιημένο σύστημα ελέγχεται και εφαρμόζεται στην προσομοίωση πεπτιδίων και αλυσίδων αμινοξέων με *ab-initio* μεθόδους.



ΚΕΦΑΛΑΙΟ 1

Θεωρία Κβαντικής Προσομοίωσης



1.1. Θεμελιώδης θεωρία Μοριακής Κβαντικής Προσομοίωσης

Ο βασικός σκοπός των υπολογιστικών προσομοιώσεων με την θεωρία της Μοριακής Κβαντικής Μηχανικής⁴³ αφορά την λύση της μη-σχετικιστικής χρονικά ανεξάρτητης εξίσωσης του Schrodinger, και ειδικότερα τον υπολογισμό της ηλεκτρονικής δομής των μορίων. Στην ενότητα αυτή γίνεται η παρουσίαση υπολογιστικών μεθόδων που χρησιμοποιούνται για να λύσουν την εξίσωση του Schrodinger για τα άτομα των μορίων. Περιγράφονται οι εξισώσεις που αποτελούν την βάση των υπολογισμών. Παρουσιάζονται επίσης και κάποιες προσεγγίσεις που γίνονται έτσι ώστε οι υπολογισμοί να είναι υλοποιήσιμη. Όλες αυτές οι μέθοδοι αναπτύχθηκαν ώστε να μπορούν να προγραμματιστούν σε ηλεκτρονικούς υπολογιστές.

Η μη σχετικιστική, χρονικά ανεξάρτητη, εξίσωση του Schrodinger δίνεται από τη σχέση:

$$H\Psi(R, r) = E(R)\Psi(R, r) \quad (1.1)$$

όπου

- H είναι ο Χαμιλτώνιος τελεστής του συστήματος (σε a.u).

$$H = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{i>j}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} \quad (1.2)$$

- Ψ είναι η κυματοσυνάρτηση
- $E(R)$ είναι μία συνάρτηση η οποία περιγράφει την ενέργεια του μορίου
- ∇^2 είναι ο Λαπλασιανός τελεστής (ανάδελτα-τετράγωνο), η οποία σε τρισσορθογώνιο σύστημα
- αξόνων εκφράζεται ως

$$\nabla^2 = \left(\frac{\partial^2}{\partial x^2} \right) + \left(\frac{\partial^2}{\partial y^2} \right) + \left(\frac{\partial^2}{\partial z^2} \right) \quad (1.3)$$

- m_e είναι η μάζα του ηλεκτρονίου
- m_A είναι η μάζα του πυρήνα A
- r_{iA} είναι η απόσταση του i ηλεκτρονίου με τον πυρήνα A
- r_{ij} είναι η απόσταση του i και του j ηλεκτρονίου
- R_{AB} είναι η απόσταση μεταξύ των πυρήνων A και B
- Z_A και Z_B είναι οι ατομικοί αριθμοί των πυρήνων A και B
- N και M είναι οι αριθμοί των ηλεκτρονίων και των πυρήνων αντίστοιχα.

Στην εξίσωση (1.2) ο πρώτος όρος, μετά την ισότητα, είναι ο τελεστής της κινητικής ενέργειας των ηλεκτρονίων, ο δεύτερος όρος είναι ο τελεστής της κινητικής ενέργειας



των πυρήνων, ο τρίτος όρος εκπροσωπεί τις έλξεις coulomb μεταξύ των ηλεκτρονίων και των πυρήνων, ενώ ο τέταρτος και ο πέμπτος όρος συνδέονται με τις απώσεις μεταξύ των ηλεκτρονίων και των πυρήνων αντίστοιχα⁴⁴.

Όλοι οι υπολογισμοί και προσεγγίσεις που γίνονται βασίζονται στην προσέγγιση Born-Oppenheimer, η οποία εστιάζεται στη λύση της ηλεκτρονιακής εξίσωσης (1.1) του Schrodinger για σταθερές θέσεις των πυρήνων.

1.1.1. Η προσέγγιση Born-Oppenheimer

Είναι γνωστό ότι ένα μόριο έχει δύο ειδών ενέργειες: τη μεταφορική (translation energy) και την εσωτερική ενέργεια U^{45} . Γενικά σε ένα μόριο παρατηρούνται οι εξής κινήσεις των σωματιδίων του (πυρήνες - ηλεκτρόνια):

- κίνηση των ηλεκτρονίων του μορίου
- δονητικές κινήσεις των πυρήνων
- περιστροφικές κινήσεις του μορίου

Η εσωτερική ενέργεια εκφράζεται ως άθροισμα της ηλεκτρονικής (E_e), της δονητικής (E_v) και της περιστροφικής ενέργειας (E_r)

$$U \approx E_e + E_v + E_r \quad (1.4)$$

Από τα τρία είδη εσωτερικής ενέργειας, αυτή που έχει ιδιαίτερη σημασία για τη μελέτη της ηλεκτρονικής δομής των μορίων είναι η E_e .

Το 1927 οι Born και Oppenheimer πρότειναν τον διαχωρισμό της ηλεκτρονικής από τις άλλες δύο μορφές εσωτερικής ενέργειας⁴⁶. Συγκεκριμένα παρατήρησαν ότι η μάζα του ηλεκτρονίου είναι πολύ μικρότερη από τη μάζα ενός πρωτονίου (1/1850 φορές μικρότερη), οπότε μπορεί να γίνει «διαχωρισμός» μεταξύ των ηλεκτρονίων και των κινήσεων των πυρήνων. Έτσι έδειξαν ότι η μοριακή κυματοσυνάρτηση, μπορεί να προσεγγισθεί επαρκώς από τη σχέση (1.5) εάν $(m_e/m_N)^{1/4} \ll 1$. Πρότειναν λοιπόν, την προσέγγιση που έγινε γνωστή ως προσέγγιση Born-Oppenheimer. Έστω $\Psi(\{r_i\}; \{R_A\})$ η κυματοσυνάρτηση που περιγράφει την κίνηση των ηλεκτρονίων και των πυρήνων, η οποία εκφράζεται με την παρακάτω σχέση:

$$\Psi(\{r_i\}; \{R_A\}) = \Phi_e(\{r_i\}; \{R_A\}) \Phi_{nuc}(\{R_A\}) \quad (1.5)$$

όπου

- Φ_e είναι η ηλεκτρονιακή κυματοσυνάρτηση, η οποία εξαρτάται από τις ηλεκτρονιακές συντεταγμένες και παραμετρικά από τις πυρηνικές

$$\Phi_e = \Phi_e(\{r_i\}; \{R_A\}) \quad (1.6)$$



- Φ_{nuc} είναι η πυρηνική κυματοσυνάρτηση

$$\Phi_{nuc} = \Phi_{nuc}(\{R_A\}) \quad (1.7)$$

Τότε σύμφωνα με την προαναφερθείσα προσέγγιση, αυτά τα δύο σύνολα κινήσεων, μπορούν να περιγραφούν ανεξάρτητα. Έτσι προκύπτουν οι εξισώσεις (1.8) και (1.9)

$$H_{el}\Phi_{el} = E_{el}\Phi_{el} \quad (1.8)$$

$$H_{nuc}\Phi_{nuc} = E\Phi_{nuc} \quad (1.9)$$

H_{el} είναι ο ηλεκτρονιακός Χαμιλτώνιος τελεστής

$$H_{el} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{i>j}^N \frac{1}{r_{ij}} \quad (1.10)$$

E_{el} είναι η ηλεκτρονιακή ενέργεια

H_{nuc} είναι ο πυρηνικός Χαμιλτώνιος τελεστής

$$H_{nuc} = \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + E_{tot}(\{R_A\}) \quad (1.11)$$

όπου με $E_{tot}(\{R_A\})$ είναι μια συνάρτηση που εκφράζει το δυναμικό εντός του οποίου κινούνται οι πυρήνες.

Οι λύσεις της πυρηνικής εξίσωσης του Schrodinger (1.9) παρέχουν πληροφορίες για τη δονητική, περιστροφική και μεταφορική κίνηση του μορίου. Συνεπώς η ενέργεια που δίνεται από την εξίσωση (1.9), η οποία είναι η Born-Oppenheimer προσέγγιση της ολικής ενέργειας, περιλαμβάνει την ηλεκτρονιακή, δονητική και μεταφορική ενέργεια του μορίου.

Σημειώνεται ότι η προσέγγιση Born-Oppenheimer δίνει ικανοποιητικά αποτελέσματα στις περισσότερες περιπτώσεις. Υπάρχουν όμως και εξαιρέσεις όπου η προσέγγιση αυτή δε δίνει επαρκείς λύσεις (π.χ. όταν οι ταχύτητες των πυρήνων είναι πολύ μεγάλες). Συνέπεια της προσέγγισης Born-Oppenheimer είναι ότι προκύπτει ένα μικρό σφάλμα στις ιδιότητες της βασικής κατάστασης (θεμελιώδης κατάσταση). Το σφάλμα αυτό είναι μεγαλύτερο στις ιδιότητες των διεγερμένων καταστάσεων, αλλά πάντως είναι μικρότερο από το σφάλμα που οφείλεται στις προσεγγίσεις που συνήθως υιοθετούνται για να λυθεί η εξίσωση του Schrodinger για ένα πολυατομικό μόριο.

Δύο είναι οι βασικές προσεγγίσεις για τη λύση της εξίσωσης (1.8).

1. **Εξ-υπαρχής υπολογισμοί (ab-initio).** Στους υπολογισμούς αυτούς επιλέγεται ένα πρότυπο για τις κυματοσυναρτήσεις των ηλεκτρονίων, και η εξίσωση



(1.8) επιλύεται χρησιμοποιώντας ως είσοδο δεδομένων, μόνο τιμές θεμελιωδών σταθερών και τους ατομικούς αριθμούς των πυρήνων.

2. **Ημι-εμπειρικές μέθοδοι (semi-empirical).** Στους υπολογισμούς αυτούς γίνεται χρήση απλοποιημένων μορφών της χαμιλτωνιανής, και προσαρμοσμένων παραμέτρων με τιμές που λαμβάνονται από πειραματικά δεδομένα.

Με τη χρήση των παραπάνω υπολογισμών, βρίσκεται αρχικά η ηλεκτρονική κυματοσυνάρτηση και η ηλεκτρονική ενέργεια. Στην συνέχεια μπορούν να υπολογισθούν και άλλες ενδιαφέρουσες ποσότητες όπως είναι η δομή του μορίου, μήκη δεσμών κλπ. Τέτοιοι υπολογισμοί μπορεί να είναι περιορισμένης ακρίβειας ή πολύ υψηλής ακρίβειας, ανάλογα με το μέγεθος του μοριακού συστήματος και την διαθέσιμη υπολογιστική ισχύ.

1.1.2. Η μέθοδος Hartree-Fock και η θεωρία SCF (Self Consistent Field)

Για την καλύτερη περιγραφή των ατόμων, των ενεργειών τους και των κατανομών των ηλεκτρονίων τους πρέπει να βρεθούν αριθμητικοί τρόποι επίλυσης της εξίσωσης Schrodinger διότι δεν υπάρχουν αναλυτικές λύσεις.

Η αρχή προς αυτή την κατεύθυνση έγινε από τον Hartree και η εισαχθείσα μέθοδος είναι γνωστή ως μέθοδος του αυτοσυνεπούς πεδίου (self-consistent field, SCF). Κατόπιν η τεχνική βελτιώθηκε από τους Fock και Slater ώστε να συμπεριλάβει την ενέργεια αλληλεπίδρασης των ηλεκτρονίων. Η μέθοδος SCF είναι επίσης γνωστή ως Hartree-Fock ή ως single-determinant theory.

Τα τροχιακά τα οποία επιτυγχάνονται με τη μέθοδο αυτή ονομάζονται ατομικά τροχιακά αυτοσυνεπούς πεδίου κατά Hartree-Fock (Hartree-Fock self-consistent atomic orbitals, HF-SCF-AOs) ή απλώς τροχιακά Hartree-Fock.

Η SCF μεθοδολογία βασίζεται στην υπόθεση ότι οποιοδήποτε ηλεκτρόνιο κινείται εντός δυναμικού το οποίο είναι το μέσο σφαιρικό δυναμικό που οφείλεται στα υπόλοιπα ηλεκτρόνια. Η εξίσωση Schrodinger ολοκληρώνεται τότε αριθμητικώς ως προς αυτό το ηλεκτρόνιο και το μέσο αυτό δυναμικό.

Η διαδικασία αυτή υποθέτει ότι οι κυματοσυναρτήσεις όλων των άλλων ηλεκτρονίων είναι ήδη γνωστές ώστε να είναι εφικτός ο υπολογισμός της μέσης τιμής του δυναμικού τους.



Αυτό όμως δεν συμβαίνει, δηλαδή οι κυματοσυναρτήσεις των υπολοίπων ηλεκτρονίων είναι άγνωστες και ο υπολογισμός αρχίζει υποθέτοντας τη μορφή των κυματοσυναρτήσεών τους (π.χ. υποθέτοντας ότι είναι τροχιακά τύπου Slater).

Λύνεται τότε η εξίσωση Schrodinger του συγκεκριμένου ηλεκτρονίου και η κυματοσυνάρτησή του χρησιμοποιείται στον υπολογισμό του δυναμικού που ασκείται πάνω σε κάποιο άλλο ηλεκτρόνιο.

Αυτό επαναλαμβάνεται ως προς όλα τα ηλεκτρόνια του ατόμου και το δυναμικό που προκύπτει ασκείται στο πρώτο ηλεκτρόνιο και επαναυπολογίζεται.

Συνήθως το βελτιωμένο αυτό δυναμικό διαφέρει από την αρχική υπόθεση και όλος ο κύκλος επαναλαμβάνεται έως ότου οι λύσεις όλων των ηλεκτρονίων παραμένουν αμετάβλητες σε έναν τελευταίο κύκλο υπολογισμού. Τότε τα τροχιακά είναι αυτοσυνεπή.

Η κανονικοποιημένη HF κυματοσυνάρτηση Φ_0 , με την οποία περιγράφεται η θεμελιώδης κατάσταση του μορίου, εκφράζεται από την n-ηλεκτρονιακή ορίζουσα Slater ως εξής⁴⁷:

$$\Phi_0 = (n!)^{-1/2} \det|\chi_a(1)\chi_b(2)\dots\chi_{c(n)}| \quad (1.12)$$

Η ηλεκτρονιακή ενέργεια της θεμελιώδους κατάστασης δίνεται από τη σχέση:

$$E = \langle \Phi_0 | H | \Phi_0 \rangle \quad (1.13)$$

όπου H είναι ο Χαμιλτώνιος τελεστής (1.10).

Αντικαθιστώντας στην εξίσωση (1.13) την (1.10) και μετά από μια σειρά πράξεων προκύπτει η σχέση:

$$E_0 = \langle \Phi_0 | H | \Phi_0 \rangle = \sum_a \langle a | h | a \rangle + \frac{1}{2} \sum_{ab}^n \langle ab | ab \rangle = \sum_a [a | h | a] + \frac{1}{2} \sum_{ab} [aa | bb] - [aa | ba] \quad (1.14)$$

Όπου

- $\int \chi_a^*(1) h(1) \chi_a(1) dx_1 = \langle a | h | a \rangle$
- $\int \chi_a^*(1) \chi_a(1) r_{12}^{-1} \chi_b^*(2) \chi_b(2) dx_1 dx_2 = [aa | bb]$
- $\int \chi_a^*(1) \chi_b(1) r_{12}^{-1} \chi_b^*(2) \chi_a(2) dx_1 dx_2 = [aa | ba]$

η οποία αποτελεί μία έκφραση της ενέργειας συναρτήσεως των τροχιακών spin.

Για την παραγωγή των εξισώσεων Hartree-Fock χρησιμοποιείται η τεχνική των συναρτησιακών μεταβολών (functional variations). Κατά τη μέθοδο⁴⁴ αυτή, εισάγεται μία δοκιμαστική συνάρτηση Φ' , στην οποία αντιστοιχεί η ενέργεια $E[\Phi']$:



$$E[\Phi'] = \langle \Phi' | H | \Phi' \rangle \quad (1.15)$$

Η Φ' μεταβάλλεται κατά μία αυθαίρετα μικρή ποσότητα :

$$\Phi' \rightarrow \Phi' + \delta\Phi' \quad (1.16)$$

οπότε η ενέργεια δίνεται από τη σχέση:

$$E = \langle \Phi' | H | \Phi' \rangle = \langle \Phi' + \delta\Phi' | H | \Phi' + \delta\Phi' \rangle = E[\Phi] + \langle \delta\Phi' | H | \Phi' \rangle + \langle \Phi' | H | \delta\Phi' \rangle + E[\Phi'] + \delta E + \dots \quad (1.17)$$

Στη μέθοδο των συναρτησιακών μεταβολών αναζητείται η Φ' για την οποία η $E[\Phi']$ είναι ελάχιστη, δηλαδή $\delta E = 0$. Επειδή τα τροχιακά spin πρέπει να είναι ορθοκανονικά θα πρέπει να ισχύει η σχέση:

$$\int \chi_i^*(1) \chi_j(1) dx_1 = \delta_{ij} \quad (1.18)$$

Λαμβάνοντας λοιπόν υπόψη τον παραπάνω περιορισμό και μετά από πράξεις προκύπτει η εξής σχέση:

$$h(1) \chi_a(1) + \sum_{b \neq a} \left[\int \chi_b(2) r_{12}^{-1} \chi_a(1) dx_2 - \sum_{b \neq a} \left[\int \chi_b^*(2) \chi_a(2) r_{12}^{-1} \right] \chi_b(1) dx_2 \right] = \varepsilon_a \chi_a(1) \quad (1.19)$$

η οποία αποτελεί μια εξίσωση Hartree-Fock.

Ο $h(1)$ δίνεται από τη σχέση

$$h(1) = -\frac{1}{2} \nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}} \quad (1.20)$$

όπου οι τελεστές μετά την ισότητα, εκπροσωπούν την κινητική ενέργεια και την έλξη ενός ηλεκτρονίου από τον πυρήνα.

Επίσης η ε_a είναι η ενέργεια του ηλεκτρονίου που περιγράφεται από το τροχιακό spin χ_a .

Ο όρος $\int \chi_b(2) r_{12}^{-1} dx_2$ στην εξίσωση (1.19) ονομάζεται *τελεστής Coulomb* ($J_b(1)$)

και εκπροσωπεί το μέσο (τοπικό) δυναμικό στο σημείο X_1 , το οποίο προκαλείται από ένα ηλεκτρόνιο στο τροχιακό spin χ_b , ενώ ο τελεστής ανταλλαγής (exchange operator) ($K_b(1)$) ορίζεται από τη σχέση:

$$K_b(1) \chi_a(1) = \left[\int \chi_b^*(2) r_{12}^{-1} \chi_a(2) \right] \chi_b(1) dx_2 \quad (1.21)$$

Έτσι λοιπόν η εξίσωση Hartree - Fock μπορεί να γραφτεί ως:

$$\left[h_1 + \sum_{b \neq a} J_b(1) - \sum_{b \neq a} K_b(1) \right] \chi_a(1) = \varepsilon_a \chi_a(1) \quad (1.22)$$

η οποία αποτελεί μία εξίσωση ιδιοτιμών.

Ακολούθως προσδιορίζεται ο τελεστής Fock (f) από τη σχέση (για κλειστή στοιβάδα)



$$f(1) = h(1) + \sum_b J_b(1) - K_b(1) \quad (1.23)$$

οπότε οι εξισώσεις Hartree – Fock γίνονται

$$f|\chi_\alpha\rangle = \varepsilon_\alpha|\chi_\alpha\rangle \quad (1.24)$$

Ο τελεστής Fock όμως εκφράζεται και ως άθροισμα

$$f(1) = h(1) + u^{HF}(1) \quad (1.25)$$

όπου $u^{HF}(1)$ είναι ένας μονο-ηλεκτρονιακός τελεστής, ο οποίος ονομάζεται δυναμικό Hartree – Fock και είναι ίσος με

$$u^{HF}(1) = \sum_b J_b(1) - K_b(1) \quad (1.26)$$

Στηριζόμενη στην αρχή των παραλλαγών (variational principle) τα “καλύτερα” τροχιακά spin είναι αυτά τα οποία ελαχιστοποιούν την ηλεκτρονιακή ενέργεια (ενέργεια θεμελιώδους κατάστασης). Για την επίλυση των εξισώσεων Hartree – Fock εισάγεται μία βάση και κατόπιν λύνεται ένα σύνολο από εξισώσεις πινάκων.

1.1.2.1. Εξισώσεις Roothaan – Επίλυση της Hartree Fock

Ο Roothaan⁴⁸ το 1951 έκανε την πρόταση να αναπαρασταθεί το ανάπτυγμα κάθε μοριακού χωρικού τροχιακού ως γραμμικός συνδυασμός ενός βασικού συνόλου συναρτήσεων χ_k . Με την καινοτομία αυτή έγινε δυνατός ο υπολογισμός των κυματοσυναρτήσεων SCF μεγάλης ακρίβειας. Στη μέθοδο LCAO (Linear Combination of Atomic Orbitals – Γραμμικός συνδυασμός των ατομικών τροχιακών) ένα μοριακό τροχιακό ψ μπορεί να εκφραστεί ως γραμμικός συνδυασμός των ατομικών τροχιακών ϕ_i ⁴⁹:

$$\psi = \sum_i c_i \phi_i \quad (1.27)$$

Οι συντελεστές c_i υπολογίζονται εφαρμόζοντας τη θεωρία των παραλλαγών. Η πιθανότητα ένα ηλεκτρόνιο το οποίο καταλαμβάνει ένα τροχιακό ψ να βρεθεί στο τροχιακό ϕ_i της βάσης, είναι ίση με $|c_i|^2$. Έτσι ένα μοριακό τροχιακό εκφράζεται ως γραμμικός συνδυασμός από ατομικά τροχιακά⁵⁰.

Ο Roothaan πρότεινε την χρησιμοποίηση ενός συνόλου από γνωστές χωρικές συναρτήσεις βάσης, με τις οποίες διαφορικές εξισώσεις που περιέχουν τα ολοκληρωμάτα της εξίσωσης (1.28), μπορούν να μετατραπούν σε ένα σύνολο από αλγεβρικές εξισώσεις και να λυθούν με τις τεχνικές των πινάκων⁵¹.



$$f(r_1)\psi_1(r_1) = \varepsilon_1\psi_1(r_1) \quad (1.28)$$

Κατόπιν εισήχθηκε ένα σύνολο από K γνωστές συναρτήσεις βάσης, οι οποίες στηρίζονται στην τεχνική LCAO:

$$\psi_i = \sum_{\mu=1}^K C_{\mu i} \Phi_{\mu} \quad i=1,2,\dots,K \quad (1.29)$$

όπου $C_{\mu i}$ είναι άγνωστοι συντελεστές. Το πρόβλημα πλέον επικεντρώνεται στον υπολογισμό των συντελεστών $C_{\mu i}$.

Όταν η επέκταση της εξίσωσης (1.29) αντικατασταθεί στην εξίσωση (1.28) προκύπτει

$$f(1) \sum_{\nu} C_{\nu i} \Phi_{\nu}(1) = \varepsilon_i \sum_{\nu} C_{\nu i} \int \Phi_{\mu}^*(1) \Phi_{\nu}(1) dr_1 \quad (1.30)$$

Στη συνέχεια πολλαπλασιάζονται και τα δύο μέλη της εξίσωσης (1.30) με τη συνάρτηση βάσης $\Phi_{\mu}^*(1)$ και με ολοκλήρωση ως προς d_1 προκύπτει η

$$\sum_{\nu} C_{\nu i} \int \Phi_{\mu}^*(1) f_1(1) \Phi_{\nu}(1) dr_1 = \varepsilon_i \sum_{\nu} C_{\nu i} \int \Phi_{\mu}^*(1) \Phi_{\nu}(1) dr_1 \quad (1.31)$$

Εισάγονται δύο πίνακες, ο πίνακας αλληλοεπικάλυψης S (overlap matrix) και ο πίνακας Fock F . Ο πίνακας S είναι ερμιτιανός τα στοιχεία του είναι:

$$S_{\mu\nu} = \int \Phi_{\mu}^*(1) \Phi_{\nu}(1) dr_1 \quad (1.32)$$

Ο S γενικά δεν είναι μοναδιαίος πίνακας, επειδή δεν είναι απαραίτητο οι συναρτήσεις βάσης να είναι ορθογώνιες. Οι τιμές των ολοκληρωμάτων αλληλοεπικάλυψης βρίσκονται στην περιοχή $0 \leq |S_{\mu\nu}| \leq 1$. Τα διαγώνια στοιχεία δίδονται από τη σχέση $S_{\mu\mu}=1$. Το πρόσημο των μη - διαγώνιων στοιχείων εξαρτάται από το σχετικό πρόσημο των δύο συναρτήσεων βάσης και από το σχετικό προσανατολισμό τους στο χώρο. Σε περίπτωση που υπάρχει γραμμική εξάρτηση τότε η τιμή του ολοκληρώματος είναι περίπου μονάδα.

Ο πίνακας Fock F , είναι επίσης ερμιτιανός, με στοιχεία

$$F_{\mu\nu} = \int \Phi_{\mu}^*(1) f(1) \Phi_{\nu}(1) dr_1 \quad (1.33)$$

Έτσι η εξίσωση (1.31) διαμορφώνεται ως εξής:

$$\sum_{\nu} F_{\mu\nu} C_{\nu i} = \varepsilon_i \sum_{\nu} S_{\mu\nu} C_{\nu i} \quad i=1,2,\dots,K \quad (1.34)$$

Η εξίσωση (1.34) αποτελεί μία έκφραση ενός συνόλου K γραμμικών ταυτόχρονων εξισώσεων (μία για κάθε τιμή του ν) με άγνωστες μεταβλητές τους συντελεστές $C_{\mu i}$.



Οι εξισώσεις αυτές είναι γνωστές ως εξισώσεις *Roothaan* και μπορούν να γραφούν με τη μορφή πινάκων ως εξής:

$$FC=SC\varepsilon \quad (1.35)$$

όπου

- C είναι ένας KxK πίνακας με στοιχεία τους συντελεστές $C_{\mu i}$ (ιδιοδιανύσματα). Ο πίνακας C ονομάζεται πίνακας ιδιοδιανυσμάτων.
- ε είναι ένας διαγώνιος KxK πίνακας των ενεργειών των τροχιακών ε_i (ιδιοτιμές).

Τα μοριακά τροχιακά μπορούν να γραφούν ως εξής:

$$\psi_i(r) = (\chi_1 \chi_2 \dots \chi_K) \begin{pmatrix} c_{i1} \\ c_{i2} \\ \vdots \\ c_{iK} \end{pmatrix} \quad (1.36)$$

1.1.2.2. Αξιολόγηση Της Μεθόδου Hartree- Fock

Η ενέργεια HF, στις περισσότερες των περιπτώσεων (άνω του 98%) αντιπροσωπεύει την ακριβή ιδιοτιμή, της μη σχετικιστικής χαμιλτωνειανής του μοριακού συστήματος. Συνήθως, όμως ενδιαφέρον δεν παρουσιάζουν οι απόλυτες τιμές της ηλεκτρονικής ενέργειας αλλά οι σχετικές ενεργειακές διαφορές τους. Για ένα συγκεκριμένο σύστημα, ενδιαφέρον εμφανίζει η μελέτη της ενεργειακής διαφοράς $\Delta E = E_A - E_B$ ανάμεσα σε δύο καταστάσεις A και B του συστήματος, καθώς το A μεταβαίνει στο B: $A \rightarrow B$.

1.1.2.3. Σφάλματα της Μεθόδου HF

Σε πολλές περιπτώσεις η μέθοδος HF, έχει αποδειχτεί ότι δίνει καλά αποτελέσματα, τα οποία είναι σε συμφωνία με τα πειραματικά δεδομένα, με σφάλμα της τάξης 1 kcal/mol όσων αφορά ενέργειες οι οποίες αναφέρονται σε μεταβολές των διαμορφώσεων (φράγμα αναστροφής, cis-trans ισομερισμός κ.λ.π.) ενέργειες δεσμών υδρογόνου, ενέργειες φωτονίων, ενέργειες ενυδάτωσης κ.λ.π.

Το κοινό χαρακτηριστικό σε όλες τις περιπτώσεις εφαρμογής της μεθόδου HF είναι ότι δεν υπάρχει ουσιαστική ηλεκτρονική επαναδιευθέτηση κατά την αλλαγή του συστήματος από μία κατάσταση A σε μία κατάσταση B. Αυτό σημαίνει ότι δεν



δημιουργείται ή διασπάται χημικός δεσμός, και ότι τα μοριακά τροχιακά τροποποιούνται σε μικρότερο βαθμό. Αν αυτές οι συνθήκες ικανοποιούνται, το σφάλμα το οποίο συνδέεται με την προσέγγιση HF, συνήθως παραμένει σταθερό καθ' όλη τη διάρκεια των δομικών αλλαγών.

1.1.3. Επιλογή βάσεων

Θεμελιώδους σημασίας για την όσο το δυνατόν πιο ακριβή μοριακή κβαντική προσομοίωση είναι η επιλογή της κατάλληλης βάσης και ο υπολογισμός της συμβολής της συσχέτισης των ηλεκτρονίων. Θα πρέπει να αναφερθεί ότι οι βάσεις που είναι επαρκείς (περισσότερο πλήρης) και δίνουν ακριβή αποτελέσματα συγχρόνως για την ενέργεια αλλά και για αρκετές μοριακές ιδιότητες, αποτελούνται από μεγάλο αριθμό ατομικών τροχιακών, μεταξύ των οποίων περιλαμβάνονται συναρτήσεις πόλωσης και διάχυσης. Για το λόγο αυτό οι υπολογισμοί με τέτοιες βάσεις είναι πολύ δαπανηροί ακόμα και για μικρά μόρια όπως το βενζόλιο (C_6H_6). Στην πράξη λοιπόν, σχεδιάζεται η βάση ώστε αφ' ενός να απαιτεί μικρό υπολογιστικό κόστος και αφ' ετέρου να δίνει επαρκή και φυσικώς αποδεκτά αποτελέσματα για τις ιδιότητες που αποτελούν αντικείμενο μελέτης.

1.1.3.1. Κατηγορίες συνόλων βάσεων (basis sets)

Για τη διεξαγωγή των μοριακών κβαντικών υπολογισμών χρησιμοποιούνται συνήθως δύο κατηγορίες ατομικών τροχιακών που συνιστούν βάσεις: α) τα τροχιακά τύπου Slater (Slater Type Orbitals - STO) και β) τα τροχιακά τύπου Gaussian (Gaussian Type Orbitals - GTO). Για λόγους που θα αναφερθούν ακολούθως, η συντριπτική πλειοψηφία των μοριακών υπολογισμών που υπάρχουν στη βιβλιογραφία την τελευταία δεκαετία έχουν διεξαχθεί με βάσεις ατομικών τροχιακών τύπου GTO. Για να γίνουν σαφή όμως τα πλεονεκτήματα των GTO και οι διαφορές τους με τα STO, στις παραγράφους που ακολουθούν γίνεται συνοπτική παρουσίαση τόσο των GTO όσο και των STO.

1.1.3.1.1. Τροχιακά τύπου Slater (STO)

Το 1930 ο Slater⁵² πρότεινε την εκθετική συνάρτηση

$$\chi_{nlm}(r, \theta, \phi) = N_n r^{n-1} \exp(-\zeta r) Y_{lm}(\theta, \phi)$$

η οποία καλείται τροχιακό Slater (STO).



Στην παραπάνω σχέση εμφανίζονται οι εξής όροι:

- N_n είναι η σταθερά κανονικοποίησης
- με $Y_{lm}(\theta, \varphi)$ είναι οι σφαιρικές αρμονικές
- n, l, m είναι οι κβαντικοί αριθμοί
- ζ είναι μία παράμετρος η οποία χρησιμοποιείται για την εύρεση των βέλτιστων τιμών των συντελεστών.

Κάθε ατομικό τροχιακό AO, αποτελείται από το γραμμικό συνδυασμό ενός ή περισσότερων τροχιακών τύπου Slater (STO). Κάθε μοριακό τροχιακό MO είναι ο γραμμικός συνδυασμός ατομικών τροχιακών τύπου Slater (LC-STO MO, Linear Combination - Slater Type Orbital). Στα πολυατομικά μόρια η μέθοδος LC-STO περιλαμβάνει ένα ή περισσότερα STO εστιασμένα στο κέντρο του κάθε ατόμου. Η παρουσία περισσότερων των δύο ατόμων προκαλεί δυσκολίες στον υπολογισμό των απαιτούμενων ολοκληρωμάτων. Σε ένα τριατομικό μόριο τα ολοκληρώματα που πρέπει να επιλυθούν είναι τριών, δύο και ενός κέντρου.

Έστω $\varphi_1, \varphi_2, \dots, \varphi_n$ είναι τα MO τροχιακά χώρου των ηλεκτρονίων 1, 2, ..., n, τότε η χαμιλτωνιανή θα περιλαμβάνει μόνον όρους ενός ή δύο ηλεκτρονίων.

Ένα τυπικό διηλεκτρονιακό ολοκλήρωμα είναι το εξής:

$$\iint \dots \int \varphi_1^* \varphi_2^* \dots \varphi_n^* \frac{1}{r_{12}} \varphi_1 \varphi_2 \dots \varphi_n du_1 du_2 \dots du_n = \iint |\varphi_1|^2 |\varphi_2|^2 \frac{1}{r_{12}} du_1 du_2 \quad (1.37)$$

Κάθε MO μπορεί να αναπτυχθεί σε όρους ενός συνόλου μονοηλεκτρονιακών συναρτήσεων $\chi_1, \chi_2, \dots, \chi_n$

$$\varphi_1 = c_1 \chi_1(1) + c_2 \chi_2(1) + \dots + c_m \chi_m(1) \quad (1.38a)$$

$$\varphi_2 = d_1 \chi_1(1) + d_2 \chi_2(1) + \dots + d_m \chi_m(1) \quad (1.38b)$$

όπου c και d είναι οριζόμενοι συντελεστές.

Υποκαθιστώντας τις σχέσεις αυτές στο ολοκλήρωμα της εξίσωσης (1.37) προκύπτει ένα άθροισμα ολοκληρωμάτων στο οποίο μετέχουν οι συναρτήσεις που ανήκουν σε μία βάση. Ένα τέτοιο τυπικό ολοκλήρωμα είναι το παρακάτω

$$c_i^* c_j d_k^* d_l \iint \chi_i(1) \chi_k^*(2) \chi_j(2) \frac{1}{r_{12}} du_1 du_2 \quad (1.39)$$

Εάν οι συναρτήσεις $\chi_i, \chi_j, \chi_k, \chi_l$ είναι επικεντρωμένες σε τέσσερις διαφορετικούς πυρήνες τότε η εξίσωση (1.39) είναι ένα ολοκλήρωμα τεσσάρων κέντρων.



Παίρνοντας m διαφορετικές επιλογές για κάθε συναρτησιακό σύνολο της εξίσωσης (1.40) αναμένεται ότι ο αριθμός των ολοκληρωμάτων άπωσης των ηλεκτρονίων στη βάση θα αυξηθεί κατά m^4

Μία λεπτομερής ανάλυση η οποία λαμβάνει υπόψη τη μεταθετική συμμετρία δείχνει ότι στο συναρτησιακό σύνολο πρέπει να υπολογιστούν $(m^4+2m^3+3m^2+2m)/8$ ολοκληρώματα δύο ηλεκτρονίων. Για ακριβείς υπολογισμούς SCF σε τυπικά μόρια μικρού και μεσαίου μεγέθους, πρέπει να χρησιμοποιηθεί σύνολο συναρτήσεων με 20 έως και 100 STO τα οποία παράγουν 20.000 έως και 13.000.000 διηλεκτρονικά ολοκληρώματα. Υπολογισμοί ολοκληρωμάτων τριών και τεσσάρων κέντρων του STO βασικού συνόλου είναι χρονοβόρα και αυτό οδήγησε σε δύο άλλους τρόπους επιλογής βασικού συνόλου.

1.1.3.1.2. Τροχιακά τύπου Gaussian (GTO)

Για την απλοποίηση των υπολογισμών των ολοκληρωμάτων το 1950 ο Boys⁵³ πρότεινε τη χρήση των τροχιακών τύπου Gaussian (GTO) αντί των STO στην LC-AO εξίσωση (1.27). Τα τροχιακά τύπου Gaussian δίνονται σε σφαιρική και σε καρτεσιανή μορφή :

$$\chi_{nlm}(r, \theta, \phi, \zeta) = N_{\alpha\zeta} r^{n-1} \exp(-\zeta r^2) Y_{-lm}(\theta, \phi)$$

(σφαιρική μορφή)

$$\chi_{nlm}(x, y, z, \zeta) = N_n x^l y^m z^n \exp(-\zeta r^2)$$

(καρτεσιανή μορφή)

Η ουσιαστική διαφορά μεταξύ των συναρτήσεων Gaussian και Slater επικεντρώνεται στη διαφορετική τάξη του εκθετικού όρου $\exp(-\zeta r^2)$ και $\exp(-\zeta r)$ αντίστοιχα. Οι συναρτήσεις τύπου GTO είναι σε αντίθεση με τις υδρογονοειδείς και τις τύπου Slater συναρτήσεις, στις οποίες ο εκθετικός παράγοντας είναι $\exp(-\zeta r)$. Η απλοποίηση των υπολογισμών των ολοκληρωμάτων οφείλεται στην εξής ιδιότητα των GTO, η οποία είναι γνωστή ως θεώρημα του *γινόμενου Gaussian*: «Το γινόμενο δύο GTO που έχουν διαφορετικά κέντρα, αποτελεί επίσης ένα GTO με κέντρο κάπου πάνω στη γραμμή που ενώνει τα δύο κέντρα».

Όπως και με τα τροχιακά Slater, έτσι και στα τροχιακά Gaussian χρησιμοποιείται η πραγματική μορφή των αρμονικών σφαιρικών στους μοριακούς υπολογισμούς. Κάθε s λοιπόν AO (είτε 1s, είτε 2s,...) αντιπροσωπεύεται ως γραμμικός συνδυασμός μερικών Gaussian, με διαφορετικούς τους εκθέτες των τροχιακών επειδή κάθε



Gaussian έχει την μορφή $\exp(-\zeta r^2 \alpha)$. Κάθε ατομικό τροχιακό αντιπροσωπεύεται από ένα γραμμικό συνδυασμό των Gaussian που το κάθε ένα έχει τη μορφή του τύπου $[\exp(-\zeta r^2 \alpha)]$ και ούτω καθ' εξής.

1.1.3.2. Διαφορές LC-GTO SCFMO και LC-STO SCFMO

Μία συνάρτηση Gaussian δεν έχει την επιθυμητή τιμή στην περιοχή κοντά στον πυρήνα και αυτό επειδή, για την ακριβή περιγραφή της ΑΟ, θα πρέπει να εισαχθούν γραμμικοί συνδυασμοί πολλών Gaussian. Έτσι ένας υπολογισμός LC-GTO SCF MO εμπεριέχει υπολογισμό περισσότερων ολοκληρωμάτων από ότι ένας αντίστοιχος υπολογισμός LC-GTO SCF MO, διότι ο αριθμός των διηλεκτρονιακών ολοκληρωμάτων είναι ανάλογος με την τέταρτη δύναμη του βασικού συνόλου. Όμως, ο υπολογισμός των ολοκληρωμάτων Gaussian απαιτεί λιγότερο χρόνο από ότι τα ολοκληρώματα Slater. Αυτό οφείλεται στο ότι, το γινόμενο δύο συναρτήσεων Gauss, επικεντρωμένο σε δύο διαφορετικά σημεία είναι ισοδύναμο με ένα απλό Gaussian επικεντρωμένο σε ένα τρίτο σημείο. Έτσι όλα τα διηλεκτρονιακά ολοκληρώματα τριών και τεσσάρων κέντρων μπορούν να απλοποιηθούν σε ολοκληρώματα δύο κέντρων.

1.1.3.3. Απλοποιήσεις των GTO

Το συνεπτυγμένο σύνολο Gaussian είναι ένας διαχωρισμός των συναρτησιακών συνόλων GTO ο οποίος χρησιμοποιείται συχνά. Σε αυτό το συνεπτυγμένο σύνολο Gaussian, δεν μεταβάλλονται ανεξάρτητα όλοι οι συντελεστές των συναρτησιακών συνόλων, αλλά ο λόγος των συντελεστών μεταξύ ορισμένων συναρτήσεων βάσης παραμένει σταθερός σε προκαθορισμένες τιμές. Αυτό έχει ως αποτέλεσμα τη μείωση του υπολογιστικού χρόνου και μικρή απώλεια από την ακρίβεια των υπολογισμών.

1.1.3.4. Ελάχιστο Βασικό Σύνολο

Ένα ελάχιστο βασικό σύνολο (minimal basis set) αποτελείται από ένα STO για κάθε εσωτερικό και δεσμικό – φλοιό του ΑΟ του κάθε ατόμου. Για παράδειγμα στο μόριο C_2H_2 , ένα ελάχιστο σύνολο συναρτήσεων (βάσης) αποτελείται από STO, 1s, 2s, 2p_x, 2p_y, 2p_z, για κάθε ΑΟ άνθρακα, και από ένα 1s STO για κάθε ΑΟ υδρογόνου. Υπάρχουν δηλαδή πέντε STO σε κάθε άνθρακα και ένα σε κάθε υδρογόνο. Το παραγόμενο σύνολο τότε αποτελείται από 12 συναρτήσεις βάσης.



Αυτό το σύνολο περιέχει:

α. σε κάθε άνθρακα, δύο s-τύπου STO και ένα p-τύπου STO, και

β. σε κάθε υδρογόνο, ένα s-τύπου STO.

Ένα τέτοιο σύνολο συμβολίζεται ως (2s 1p) για τις συναρτήσεις του C και 1s για τις συναρτήσεις του H. Πιο συνεπτυγμένος ακόμη συμβολισμός είναι (2s1p/1s).

1.1.3.5. Βάση Διπλού Ζήτα (DZ)

Η βάση διπλού ζήτα λαμβάνεται με αντικατάσταση κάθε STO του ελάχιστου συναρτησιακού συνόλου από δύο STO τα οποία διαφέρουν μεταξύ τους στον εκθέτη του τροχιακού ζ. Ένα απλό STO, δεν είναι μία ακριβής παρουσίαση ενός ΑΟ ενώ η χρήση δύο STO δίνει μία ουσιαστική βελτίωση. Για παράδειγμα, στην ένωση C_2H_2 , μία βάση διπλού-ζήτα αποτελείται από δύο 1s STO σε κάθε H, και σε κάθε άνθρακα από δύο 1s STO, δύο 2s STO, δύο $2p_x$, δύο $2p_y$ και δύο $2p_z$ STO. Το σύνολο που παράγεται αποτελείται από 24 συναρτήσεις. Αυτό είναι μία βάση (4s 2p/2s).

1.1.3.6. Σύνολα βάσεων Split-Valence

Ένα, βασικό σύνολο split valence χρησιμοποιεί δύο STO για κάθε δεσμικό φλοιό του ΑΟ, αλλά μόνο ένα STO για κάθε εσωτερικό φλοιό του ΑΟ. Το ΑΟ πολώνεται αφού σχηματισθεί το μόριο. Για να περιγραφεί αυτή η πόλωση, πρέπει να εισαχθούν συναρτήσεις STO, που οι κβαντικοί τους αριθμοί /να είναι μεγαλύτεροι από το μέγιστο / στο αντίστοιχο ελεύθερο άτομο. Για παράδειγμα, για έναν υδρογονάνθρακα, θα πρέπει να εισαχθούν στη βάση διπλού ζήτα οι ακόλουθες συναρτήσεις πολωσιμότητας $2p_x$, $2p_y$, $2p_z$ για κάθε υδρογόνο και ένα σύνολο συναρτήσεων 5d σε κάθε άτομο άνθρακα. Αυτό δίνει μία (4s 2p 1d/ 2s 1p) συναρτησιακό σύνολο διπλού-ζήτα και πολωσιμότητας. Μπορούν επίσης να εισαχθούν και μεγαλύτερες συναρτήσεις πολωσιμότητας.

Στους μοριακούς υπολογισμούς συχνά γίνεται η χρήση συνεπτυγμένης βάσης που αντικαθιστά κάθε STO. Ένα συνεπτυγμένο Gaussian (Συν. G) έχει τη μορφή $\sum_i a_i f_i$, όπου f_i είναι Gaussian συναρτήσεις οι οποίες καλούνται και πρωταρχικά Gaussian - primitive Gaussian. Οι συντελεστές επιλέγονται έτσι ώστε να δίνουν καλή προσαρμογή στο STO και παραμένουν σταθεροί σε ένα μοριακό υπολογισμό. Ο Pople και οι συνεργάτες του έκαναν εκτενείς υπολογισμούς SCF με αυτήν την προσέγγιση, χρησιμοποιώντας το βασικό σύνολο STO-3G, 4-31G και 6-31G*.



1.1.3.6.1. Βάση τύπου STO-NG

Μία βάση STO-NG έχει τα ακόλουθα χαρακτηριστικά :

- χρησιμοποιεί ένα ελάχιστο σύνολο συναρτήσεων ενός STO για κάθε ΑΟ.
- Οι εκθέτες των τροχιακών έχουν σταθερές τιμές, οι οποίες βρέθηκαν από υπολογισμούς σε μικρά μόρια.
- Για να αποφευχθούν ολοκληρώματα που η επίλυση τους είναι χρονοβόρα, κάθε STO προσεγγίζεται ως γραμμικός συνδυασμός των N Gaussian συναρτήσεων. Οι συντελεστές των N Gaussian συναρτήσεων είναι γραμμικοί συνδυασμοί και οι Gaussian εκθέτες των τροχιακών επιλέγονται έτσι ώστε να δίνουν την καλύτερη προσαρμογή με τη μέθοδο των ελαχίστων τετραγώνων.

1.1.3.6.2. Βάση τύπου STO-3G

Στις βάσεις STO-NG συνήθως το $N=3$. Αφού ένας γραμμικός συνδυασμός τριών Gaussian είναι μόνο μία προσέγγιση ενός STO, η STO-3G μέθοδος δίνει λιγότερο καλά αποτελέσματα από έναν υπολογισμό STO με ελάχιστο βασικό σύνολο. Όπως αναφέραμε, ένα ελάχιστο βασικό σύνολο ενός υδρογονάνθρακα συμβολίζεται ως ($2s\ 1p/1s$). Αφού κάθε STO αντικαθίσταται από ένα γραμμικό συνδυασμό τριών πρωταρχικών Gaussian (που είναι ένα συνεπτυγμένο Gaussian) τότε η βάση STO-3G για το υδρογονάνθρακα θα συμβολίζεται ως ($6s3p/3s$) και με συνεπτυγμένο Gaussian συμβολίζεται ως [$2s1p/1s$]. Η παρένθεση δηλώνει το πρωταρχικό Gaussian, ενώ οι αγκύλες το συνεπτυγμένο σύνολο βάσης.

1.1.3.6.3. Βάση τύπου 4-31G

Η βάση 4-31G είναι ένα split-valence σύνολο συναρτήσεων που χρησιμοποιεί συνεπτυγμένα Gaussian. Κάθε εσωτερικός φλοιός $1s$ του ΑΟ λαμβάνεται ως γραμμικός συνδυασμός τεσσάρων Gaussian με καθορισμένους συντελεστές. Για κάθε δεσμικό φλοιό $2s$ ή $2p$ ΑΟ, χρησιμοποιείται η διπλού-ζήτα προσέγγιση, παίρνοντας μία $2s$ (ή $2p$) συνάρτηση, η οποία είναι ένας καθορισμένος γραμμικός συνδυασμός τριών Gaussian και μία δεύτερη $2s$ (ή $2p$) συνάρτηση η οποία αποτελείται από μία απλή Gaussian με εκθέτη διαφορετικό από αυτό των τριών Gaussian του γραμμικού συνδυασμού. Το $1s$ κάθε υδρογόνου, αντιμετωπίζεται όπως το $2s$ δεσμικό ΑΟ των C, N ή O. Η βάση 4-31G υπερέχει ελαφρά της βάσης διπλού - ζήτα.



1.1.3.6.4. Βάση τύπου 6-31G*

Η βάση 6-31G* είναι καλύτερη από την 4-31G. Στη βάση αυτή γίνεται χρήση γραμμικού συνδυασμού έξι Gaussian, για κάθε εσωτερικό φλοιό 1s του ΑΟ και επιπλέον προτίθεται ένα απλό σύνολο d-τύπου Gaussian πολωμένων συναρτήσεων για κάθε άτομο εκτός του υδρογόνου.

1.1.3.7. Slater ή Gaussian

Η επιλογή των Slater ή Gaussian συναρτησιακών συνόλων εξαρτάται από το είδος του υπολογισμού. Για υπολογισμούς ατόμων, οι συναρτήσεις Slater είναι εμφανώς καλύτερες διότι τα ατομικά (ενός κέντρου) ολοκληρώματα είναι πιο εύκολα για STO από ότι για GTO. Οι περισσότεροι SCF υπολογισμοί γραμμικών μορίων χρησιμοποιούν STO, διότι η γραμμική γεωμετρία επιτρέπει τη χρήση ειδικών αριθμητικών μεθόδων. Για μη γραμμικά μόρια, η μέθοδος GTO, είναι γενικά η καλύτερη επιλογή για υπολογισμούς SCF λόγω του πιο γρήγορου ρυθμού επίλυσης των ολοκληρωμάτων, ανεξάρτητα από το πλήθος των ολοκληρωμάτων που πρέπει να επιλυθούν.

1.1.4. Μέθοδοι υπολογισμού της ηλεκτρονικής συσχέτισης

Ο υπολογισμός της ενέργειας ηλεκτρονικής συσχέτισης είναι ιδιαίτερα σημαντικός στους υπολογισμούς προσομοίωσης ασθενών διαμοριακών αλληλεπιδράσεων. Σε όλες τις προσομοιώσεις της διατριβής αυτής συμπεριλαμβάνονται υπολογισμοί με μεθόδους ηλεκτρονικής συσχέτισης ενώ ιδιαίτερη σημασία δίνεται στο κεφάλαιο 3.

Η ηλεκτρονικής ενέργειας συσχέτισης συμβολίζεται ως E_{cor} . Η ενέργεια αυτή ορίζεται από τη σχέση:

$$E_{cor} \equiv E_{nr} - E_{HF} \quad (1.40)$$

όπου E_{nr} είναι η ακριβής ενέργεια του n-ηλεκτρονιακού, μη σχετικιστικού (non-relativistic) χαμιλτώνιου τελεστή και E_{HF} είναι η ενέργεια Hartree – Fock. Ο ορισμός αυτός είναι επαρκής υπό την προϋπόθεση ότι τα σχετικιστικά φαινόμενα είναι μικρά.

Τρεις είναι κυρίως οι μέθοδοι που χρησιμοποιούνται για τον υπολογισμό της συμβολής της συσχέτισης των ηλεκτρονίων:

- (α) η αλληλεπίδραση διαμορφώσεων (Configuration Interaction – CI)⁵⁴
- (β) η θεωρία διαταραχών (Many Body Perturbation Theory - MBPT)⁵⁵ και
- (γ) η θεωρία Coupled Cluster (CC)⁵⁶.



Η πρώτη στηρίζεται στη θεωρία των παραλλαγών, ενώ οι άλλες δύο όχι. Στις επόμενες υποενότητες γίνεται μια σύντομη παρουσίαση των μεθόδων αυτών.

1.1.4.1. Θεωρία των διαταραχών Moller- Plesset

Στη θεωρία Moller – Plesset, ο όρος των ηλεκτρονιακών απώσεων εκφράζεται ως διαταραχή στο Χαμιλτώνιο τελεστή της HF:

$$H = H_0 + V \quad (1.41)$$

όπου

$$H_0 = \sum_i^n f(i) + V_{nuc} = \sum_i^n [h(i) + u^{HF}(i)] + V_{nuc} \quad (1.42)$$

ενώ η διαταραχή V είναι ίση με

$$V = \sum_{i>j}^n r_{ij}^{-1} - \sum_i^n u^{HF}(i) \quad (1.43)$$

Η πρώτης τάξης ενέργεια δίνεται από τη σχέση

$$E^{(1)} = \langle \Psi_0 | V | \Psi_0 \rangle = -\frac{1}{2} \sum_{ij} (ij || ij) \quad (1.44)$$

όπου η HF κυματοσυνάρτηση $|\Psi_0\rangle$, είναι ιδιοσυνάρτηση του H_0

$$H_0 |\Psi_0\rangle = E^{(0)} |\Psi_0\rangle \quad (1.45)$$

Η ενέργεια Hartree-Fock E_0 δίνεται από την εξίσωση

$$E_0 = E^{(0)} + E^{(1)} \quad (1.46)$$

Οι εκφράσεις της ενέργειας μέχρι την 4η τάξη δίνονται από τις σχέσεις⁵⁷:

$$E^{(2)} = -\sum_m \frac{\langle 0 | V | D_m \rangle \langle D_m | V | 0 \rangle}{(E_m - E_0)} \quad (1.47)$$

$$E^{(3)} = \sum_{mn} \frac{\langle 0 | V | D_m \rangle \langle D_m | V | D_n \rangle \langle D_n | V | 0 \rangle}{(E_m - E_0)(E_n - E_0)} - \sum_{mn} \frac{\langle 0 | V | D_m \rangle \langle D_m | V | 0 \rangle \langle 0 | V | 0 \rangle}{(E_m - E_0)^2} \quad (1.48)$$

$$E^{(4)} = \sum_{mn} \sum_p \frac{\langle 0 | V | D_m \rangle \langle D_m | V | X_p \rangle \langle X_p | V | D_n \rangle \langle D_n | V | 0 \rangle}{(E_m - E_0)(E_p - E_0)(E_n - E_0)} + \sum_{mn} \frac{\langle 0 | V | D_m \rangle \langle D_m | V | 0 \rangle \langle 0 | V | D_n \rangle \langle D_n | V | 0 \rangle}{(E_m - E_0)(E_n - E_0)^2} \quad (1.49)$$



$$-\sum_{mn} \frac{\langle 0|V|D_m\rangle\langle D_m|V|0\rangle\langle 0|V|D_n\rangle\langle D_n|V|0\rangle}{(E_m - E_0)^3}$$

$$+ 2\sum_{mn} \frac{\langle 0|V|D_m\rangle\langle D_m|V|D_n\rangle\langle D_n|V|0\rangle\langle 0|V|0\rangle}{(E_m - E_0)^2(E_n - E_0)}$$

όπου τα m και n καλύπτουν τις διπλές διεγέρσεις και X_p περιλαμβάνει τις απλές, διπλές, τριπλές και τετραπλές διεγέρσεις (S, D, T και Q).

Η \tilde{E}_m δίνεται από τη σχέση:

$$E_m = \langle m|H_0|m\rangle \quad (1.50)$$

Η ενέργεια $E^{(4)}$ εκφράζεται επίσης και από τη σχέση⁵⁸:

$$E^{(4)} = E_R^{(4)} + E_S^{(4)} + E_D^{(4)} + E_T^{(4)} + E_Q^{(4)} \quad (1.51)$$

όπου ο όρος $E_R^{(4)}$ είναι ίσος με

$$E_R^{(4)} = -E^{(2)} \sum_n \frac{V_{0n}V_{n0}}{(E_n - E_0)} \quad (1.52)$$

και $V_{pq} = \langle p|V|q\rangle$.

Είναι γνωστό ότι η μελέτη των ιδιοτήτων σύγκλισης της σειράς MPn έχει οδηγήσει σε επικριτικές παρατηρήσεις για την απόδοσή της. Σημειώνεται ότι γενικότερα αποδεκτή είναι η θεωρία MP2.

1.1.4.2. Μέθοδοι coupled cluster (CC)

Η θεμελιώδη εξίσωση της θεωρίας CC⁵⁶ είναι η

$$\psi = e^T \Phi_0 \quad (1.53)$$

όπου ψ είναι η ακριβής μη σχετικιστική μοριακή ηλεκτρονιακή κυματοσυνάρτηση της θεμελιώδους κατάστασης, Φ_0 είναι η κανονικοποιημένη κυματοσυνάρτηση Hartree - Fock της θεμελιώδους κατάστασης και ο τελεστής e^T ορίζεται από τη σειρά Taylor:

$$e^T \equiv 1 + T + \frac{T^2}{2!} + \frac{T^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{T^k}{k!} \quad (1.54)$$

όπου T είναι ο τελεστής cluster, ο οποίος εκφράζεται με το παρακάτω άθροισμα

$$T \equiv T_1 + T_2 + \dots T_n \quad (1.55)$$

όπου n είναι ο αριθμός των ηλεκτρονίων του μορίου. Οι τελεστές διέγερσης του ενός και των δύο σωματιδίων, T_1 και T_2 αντίστοιχα ορίζονται από τις σχέσεις:



$$T_1\Phi_0 \equiv \sum_{b=n+1}^{\infty} \sum_{i=1}^n t_i^a \Phi^a, \quad T_2\Phi_0 \equiv \sum_{b=a+1}^{\infty} \sum_{\sigma=n+1}^{\infty} \sum_{j=i}^n \sum_{l=1}^{n-1} t_{ij}^{ab} \Phi_{ij}^{ab} \quad (1.56)$$

όπου Φ_i^a είναι μία ορίζουσα Slater όπου τα κατειλημμένα (occupied) τροχιακά spin u_i , έχουν αντικατασταθεί από τα εικονικά ή μη κατειλημμένα (virtual) τροχιακά spin u_a . Οι t_i^a, t_{ij}^{ab} είναι αριθμητικοί συντελεστές των οποίων η τιμή εξαρτάται από το i και το a . Ο τελεστής T_1 μετατρέπει την ορίζουσα Slater, $[u_i \dots u_n] = \Phi_0$, σε έναν γραμμικό συνδυασμό όλων των απλά διεγερμένων (singly excited) ορίζουσών Slater. Η Φ_{ij}^{ab} είναι μία ορίζουσα Slater όπου τα κατειλημμένα τροχιακά spin u_i και u_j έχουν αντικατασταθεί από τα εικονικά (κενά) τροχιακά spin u_a και u_b αντίστοιχα. Ανάλογα ορίζονται και οι υπόλοιποι τελεστές T_3, \dots, T_n . Το εξεταζόμενο σύστημα θεωρείται ότι έχει n ηλεκτρόνια. Για το λόγο αυτό δεν μπορούν να διεγερθούν περισσότερα από n ηλεκτρόνια από μια n -ηλεκτρονιακή κυματοσυνάρτηση Φ_0 . Η παρουσία του τελεστή e^T στην εξίσωση (1.54) έγινε έτσι ώστε να εκφράζεται η ψ ως γραμμικός συνδυασμός ορίζουσών Slater, οι οποίες περικλείουν τη Φ_0 και όλες τις πιθανές διεγέρσεις των ηλεκτρονίων από τα κατειλημμένα στα εικονικά τροχιακά spin. Σκοπός ενός CC υπολογισμού είναι η εύρεση των συντελεστών (amplitudes) $t_j^a, t_{ij}^{ab}, t_{ijk}^{abc}$ κ.λ.π., για όλα τα i, j, k, \dots και a, b, c, \dots

1.1.4.2.1. Μέθοδος coupled cluster doubles (CCD)

Από τη θεωρία προκύπτει ότι ο τελεστής T_2 έχει πιο σημαντική συμβολή στον T . Η προσέγγιση $T \approx T_2$ δίνει

$$\Psi_{CCD} = e^{T_2}\Phi_0 \quad (1.57)$$

Η προκύπτουσα προσέγγιση ονομάζεται coupled cluster doubles (CCD). Επειδή

$$e^{T_2} \equiv 1 + T_2 + \frac{T_2^2}{2!} \quad (1.58)$$

η κυματοσυνάρτηση Ψ_{CCD} περιέχει ορίζουσες με διπλές, τετραπλές κ.λ.π., υποκταστάσεις. Η μέθοδος CCD είναι size-consistent αλλά δεν προκύπτει από χρήση της θεωρίας των παραλλαγών (δηλ. δεν είναι variational).

1.1.4.2.2. Οι μέθοδοι coupled cluster singles, doubles και triples (CCSD και CCSD(T))



Η επόμενη προσέγγιση αξιοποιεί και τον τελεστή T_1 , δηλαδή $T = T_1 + T_2$. Η προκύπτουσα προσέγγιση είναι η CCSD (coupled cluster singles and doubles).

Η χρήση του τελεστή $T = T_1 + T_2 + T_3$ οδηγεί στη μέθοδο CCSDT⁵² (coupled cluster with singles, doubles and triples). Έχουν προταθεί διάφορες προσεγγιστικές μορφές της Ψ_{CCSDT} , εκ των οποίων η ευρύτερα χρησιμοποιούμενη είναι η Ψ_{CCSDCT} . Σ' αυτή την προσέγγιση η συμβολή των τριπλών υποκαταστάσεων έχει υπολογισθεί με τη θεωρία MP4, χρησιμοποιώντας όμως τους CCSD συντελεστές (amplitudes).

1.1.4.3 Quadratic Configuration Interaction με απλές και διπλές υποκαταστάσεις (QCISD)

Η μέθοδος προτάθηκε από τον Pople και τους συνεργάτες του⁵⁹ για τον υπολογισμό της συμβολής της συσχέτισης των ηλεκτρονίων. Σημείο εκκίνησης αποτελεί μία κυματοσυνάρτηση που προσεγγίζεται από μία ορίζουσα. Κατόπιν ορίζονται οι τελεστές απλής, διπλής και τριπλής υποκατάστασης, ως εξής:

$$T_1 = \sum_a \alpha_i^a t_i^a \quad (1.59)$$

$$T_2 = \frac{1}{4} \sum_{ijab} \alpha_{ij}^{ab} t_{ij}^{ab} \quad (1.60)$$

$$T_3 = \frac{1}{36} \sum_{ijkabc} \alpha_{ijk}^{abc} t_{ijk}^{abc} \quad (1.61)$$

όπου $t_i^a, t_{ij}^{ab}, \dots$ είναι τελεστές υποκατάστασης και οι πίνακες α περιέχουν τους συντελεστές που πρέπει να προσδιορισθούν. Οι σχέσεις (1.59 – 1.61) ορίζουν τις εξισώσεις QCISD:

$$\langle \Psi_0 | H | T_2 \Psi_0 \rangle = E^{\text{corr}} \quad (1.62)$$

$$\langle \Psi_i^a | \overline{H} | (T_1 + T_2 + T_1 T_2) \Psi_0 \rangle = \alpha_i^a E^{\text{corr}} \quad (1.63)$$

$$\langle \Psi_{ij}^{ab} | \overline{H} | (1 + T_1 + T_2 + \frac{1}{2} T_2^2) \Psi_0 \rangle = \alpha_{ij}^{ab} E^{\text{corr}} \quad (1.64)$$

Επεξηγώντας τον όρο «quadratic cent, interaction» ο Pople και οι συνεργάτες του σημείωσαν⁶⁰: "The theory is certainly related to CCSD but the approach outlined above shows that the QCISD equations can also be presented as a development of the full configuration interaction method, thereby avoiding some of the complex superstructure of coupled cluster theory". Η μέθοδος διασφαλίζει τη size consistency στην ολική ενέργεια. Ο πλήρης υπολογισμός των τριπλών υποκαταστάσεων οδηγεί



στη θεωρία QCISDT, η οποία όμως είναι ιδιαίτερα δαπανηρή, από άποψη υπολογιστικού χρόνου. Η προσέγγιση της επίδρασης των τριπλών υποκαταστάσεων με μία μέθοδο που είναι παρόμοια μ' αυτήν που χρησιμοποιείται από τη θεωρία MP4[SDTQ] οδηγεί στη μέθοδο QCISD(T).

1.1.5. Κλιμάκωση μεθόδων

Στο επόμενο πίνακα 1.1 φαίνεται η τάξη υπολογιστικής πολυπλοκότητας των μεθόδων κβαντικής προσομοίωσης ως προς το μέγεθος N του συνόλου της βάσης.

Πίνακας 1.1

Τάξη Υπολογισμού	Μέθοδος
N^4	HF
N^4	DFT (Non-Local)
N^5	MP2
N^6	MP3, CISD, MP4SDQ, CCSD, QCISD
N^7	MP4, CCSD(T), QCISD(T)
N^8	MP5, CISDT, CCSDT
N^9	MP6
N^{10}	MP7, CISDTQ, CCSDTQ

Σε κάθε υπολογισμό κβαντικής προσομοίωσης είναι πολύ σημαντική τόσο η επιλογή της βάσης όσο και της μεθόδου.



1.2. Υπολογιστικές μέθοδοι γεωμετρικής βελτιστοποίησης.

Οι μέθοδοι γεωμετρικής βελτιστοποίησης μοριακών συστημάτων χωρίζονται σε δύο κατηγορίες: στις τοπικές (local methods) και στις γενικευμένες μεθόδους (global methods), οι οποίες ερευνούν μεγαλύτερες περιοχές για την εύρεση του γενικευμένου ελαχίστου σε αντίθεση με τις τοπικές⁶¹.

1.2.1. Τοπικές μέθοδοι ελαχιστοποίησης (local methods)

Οι τοπικές μέθοδοι ελαχιστοποίησης χωρίζονται στις εξής κατηγορίες:

- Ελαχιστοποιήσεις χωρίς τη χρήση παραγώγων
- Ελαχιστοποιήσεις με τη χρήση παραγώγων

Ο διαχωρισμός αυτός γίνεται ανάλογα με το αν χρησιμοποιούνται, κατά τη διεξαγωγή της ελαχιστοποίησης, παράγωγοι ή όχι.

1.2.2. Ελαχιστοποιήσεις χωρίς τη χρήση παραγώγων

Υπάρχουν αρκετοί ελαχιστοποιητές οι οποίοι βρίσκουν τη βέλτιστη λύση μιας συνάρτησης χωρίς να χρησιμοποιούν παραγώγους. Ο πιο γνωστός αλγόριθμος της κατηγορίας αυτής είναι ο αλγόριθμος Simplex⁶².

1.2.3. Ελαχιστοποιήσεις με τη χρήση παραγώγων

Στην κατηγορία αυτή οι τοπικοί ελαχιστοποιητές βασίζονται στον υπολογισμό των παραγώγων. Στην συνέχεια αναφέρονται ορισμένες μέθοδοι τοπικής ελαχιστοποίησης οι οποίες ανήκουν στην κατηγορία αυτή.

1.2.3.1. Μέθοδοι παραγώγων

Οι μέθοδοι που βασίζονται στον υπολογισμό των παραγώγων για τον εντοπισμό των ελαχίστων χωρίζονται σε δύο κατηγορίες ανάλογα με τον τρόπο που τις υπολογίζουν.

Οι μέθοδοι αυτές διακρίνονται στις:

- 1) *Αριθμητικές.*
- 2) *Αναλυτικές μεθόδους.*

Από το 1969, όταν ο P. Pulay έγραψε το πρώτο υπολογιστικό πρόγραμμα για τον αναλυτικό προσδιορισμό των πρώτων παραγώγων των ενεργειών HF, οι μέθοδοι που χρησιμοποιούν τις παραγώγους εξελίχθηκαν σε έναν από τους πιο εντατικά μελετημένους τομείς της σύγχρονης προσομοίωσης μοριακών συστημάτων με μεθόδους κβαντικής χημείας. Οι μέθοδοι αυτές αρχικά εφαρμόστηκαν σε



υπολογισμούς HF πλήρους στοιβάδας (close shell model) και κατόπιν γενικεύτηκαν σε κυματοσυναρτήσεις, οι οποίες λαμβάνουν υπόψη τη συσχέτιση των ηλεκτρονίων.

1.2.3.2. Τοπικοί ελαχιστοποιητές με χρήση παραγώγων κατηγορίας conjugate gradient (συζύγων διευθύνσεων)

Οι μέθοδοι των συζυγών διευθύνσεων (conjugate gradient methods) είναι αρκετά οικονομικές όσον αφορά την απαιτούμενη υπολογιστική μνήμη. Στους ελαχιστοποιητές αυτής της κατηγορίας δίνεται αρχικά ένα σημείο εκκίνησης $\chi^{(1)}$ και οι πρώτες παράγωγοι της συνάρτησης $g^{(1)}$. Κατόπιν θέτεται ως πρώτη διεύθυνση αναζήτησης η $s^{(1)} = -g^{(1)}$. Η k επανάληψη αποτελείται από τα εξής βήματα, τα οποία παρατίθενται ως αλγοριθμική διαδικασία:

- 1) Εκτελείται μία γραμμική αναζήτηση κατά μήκος της διεύθυνσης $s^{(k)}$ από την οποία προκύπτει το $x^{(k+1)} = x^{(k)} + as^{(k)}$.
- 2) Ελέγχονται τα κριτήρια τερματισμού, έτσι ώστε να διαπιστωθεί εάν θα τερματίσει ή όχι η διαδικασία σύγκλισης.
- 3) Υπολογίζεται το διάνυσμα $g^{(k+1)}$.
- 4) Υπολογίζεται η παράμετρος $\beta^{(k)}$ χρησιμοποιώντας μία από τις δύο παρακάτω μεθόδους:

$$(\alpha) \quad \beta^{(k)} = \frac{g^{(k+1)T} g^{(k+1)}}{g^{(k)T} g^{(k)}} \quad \text{μέθοδος Fletcher - Reeves} \quad (1.65)$$

$$(\beta) \quad \beta^{(k)} = \frac{(g^{(k+1)} - \beta^{(k)} g^{(k)})^T g^{(k+1)}}{g^{(k)T} g^{(k)}} \quad \text{μέθοδος Polak - Ribiere} \quad (1.66)$$



1.2.3.3. Ελαχιστοποίηση με τη χρήση της μεθόδου Newton

Έστω ότι η διαδικασία ελαχιστοποίησης αποσκοπεί στην εύρεση του ελαχίστου μίας συνάρτησης $f(x)$, όπου $x \in \mathbb{R}^n$. Η μέθοδος Newton προτάθηκε για την επίλυση των N μη γραμμικών εξισώσεων που απαιτούνται για την εύρεση της ελάχιστης τιμής $g(x)=0$. Η μέθοδος Newton συγκλίνει πολύ γρήγορα. Βασικό της μειονέκτημα αποτελεί ο υπολογισμός του εσσιανού πίνακα, σε κάθε επανάληψη. Γι' αυτό το λόγο έγιναν αρκετές αλλαγές στον τρόπο σύγκλισης της από τις οποίες προέκυψαν οι αλγόριθμοι που περιγράφονται στις δύο επόμενες υποενότητες.

1.2.3.3.1. Τοπικοί ελαχιστοποιητές κατηγορίας Newton - Raphson

Στους τοπικούς ελαχιστοποιητές που χρησιμοποιούν τις παραγώγους της κατηγορίας Newton-Raphson θεωρείται ότι η αρχική συνάρτηση f εκφράζεται ως μια δυναμοσειρά δεύτερης τάξης γύρω από το σημείο x_0 .

$$f(x) \approx f(x_0) + g'(x - x_0) + \frac{1}{2}(x - x_0)' G(x - x_0) \quad (1.67)$$

Στη συνέχεια, απαιτείται η g να είναι 0, οπότε η εξίσωση (1.67) γίνεται:

$$(x - x_0) = -\frac{g}{G} - G^{-1}g \quad (1.68)$$

Εάν ο εσσιανός πίνακας (πίνακας δευτέρων παραγώγων G) είναι διαγώνιος, τότε η εξίσωση (1.68) έχει τη μορφή

$$\Delta x' = \sum_i \Delta x_i, \quad \Delta x_i = -\frac{f_i}{\epsilon_i} \quad (1.69)$$

όπου

- f_i είναι η συνιστώσα της g κατά μήκος της κατεύθυνσης του i ιδιοδιανύσματος του εσσιανού πίνακα με ιδιοτιμή ϵ_i .
- x' είναι το σύστημα συντεταγμένων

Ένα από τα προβλήματα που μπορεί να παρουσιαστεί κατά την ελαχιστοποίηση με τη χρήση της NR μεθόδου είναι το εξής: Εάν η ελαχιστοποίηση ξεκινάει από μία περιοχή όπου ο Εσσιανός πίνακας έχει αρνητικές ιδιοτιμές τότε η εξίσωση (1.69) θα δώσει κάποιο σαγματικό σημείο ή κάποιο μέγιστο. Αυτό αποτελεί ένα από τα βασικότερα προβλήματα της NR μεθόδου, το οποίο λύνεται με την προσθήκη της παραμέτρου λ στον παρονομαστή της εξίσωσης (1.69). Από την τιμή του λ εξαρτάται και το είδος του ελαχιστοποιητή που θα χρησιμοποιηθεί. Έτσι λοιπόν η εξίσωση (1.69) γίνεται:



$$\Delta x' = \sum_i \Delta x_i, \quad \Delta x'_i = -\frac{f_i}{\varepsilon_i - \lambda} \quad (1.70)$$

Οι πιο γνωστές NR μέθοδοι είναι οι: Eigenvector Following (EF), Quadratic Approximation (QA) κ.λ.π.

1.2.3.3.2. Μέθοδοι quasi – Newton

Στους ελαχιστοποιητές αυτής της κατηγορίας δίνεται αρχικά ένα σημείο εκκίνησης $x^{(1)}$, οι πρώτες παράγωγοι της συνάρτησης $g^{(1)}$ και ένας θετικά ορισμένος πίνακας των δευτέρων παραγώγων $G^{(1)}$. Στην αρχή της k επανάληψης είναι διαθέσιμα το σημείο $x^{(k)}$, η κλίση $g^{(k)}$ και μία προσέγγιση $B^{(k)}$ του εσσιανού πίνακα $G^{(k)}$. Στην επανάληψη αυτή εκτελούνται τα εξής βήματα που περιγράφουν τον αλγόριθμο:

- 1) Ελέγχονται τα κριτήρια τερματισμού, έτσι ώστε να διαπιστωθεί εάν θα τερματίσει ή όχι η διαδικασία σύγκλισης.
- 2) Επιλύεται η $B^{(k)}s^{(k)} = -g^{(k)}$ για $s^{(k)}$.
- 3) Επιλέγεται ένα «καλύτερο» σημείο $x^{(k+1)}$.
- 4) Υπολογίζεται το διάνυσμα παραγώγων (gradient) $g^{(k+1)}$.
- 5) Ενημέρωση του $B^{(k)}$ με αντικατάσταση του από τον $B^{(k+1)}$ χρησιμοποιώντας τη φόρμουλα quasi-Newton.

Δύο είναι οι πιο διαδεδομένοι τύποι για την ενημέρωση του $G^{(k)}$. Οι μέθοδοι αυτές είναι οι:

- DFP (Davidon⁶³, Fletcher και Powell⁶⁴)
- BFGS (Broyden⁶⁵, Fletcher⁶⁶, Goldfarb⁶⁷ και Shanno⁶⁸).

Χρησιμοποιώντας τους ορισμούς: $\delta^{(k)} = x^{(k+1)} - x^{(k)}$ και $\gamma^{(k)} = g^{(k+1)} - g^{(k)}$ οι τύποι που περιγράφουν την ενημέρωση στις μεθόδους BFGS και DFP είναι αντίστοιχα:

$$B^{(k+1)} = B + \frac{\gamma\gamma^T}{\delta^T\gamma} - \frac{B\delta\delta^TB}{\delta^TB\delta} \quad (\text{BFGS})$$

$$B^{(k+1)} = B + \left(1 + \frac{\delta^TB\delta}{\delta^T\gamma}\right) \frac{\gamma\gamma^T}{\delta^T\gamma} - \frac{\gamma\delta^TB + B\delta\gamma^T}{\delta^T\gamma} \quad (\text{DFP})$$



ΚΕΦΑΛΑΙΟ 2

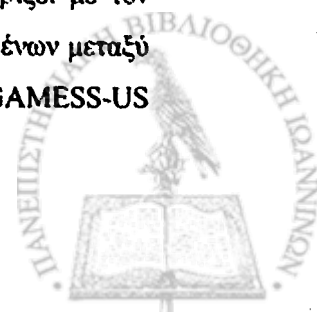
MERGAM: Ένα προγραμματιζόμενο περιβάλλον βελτιστοποίησης με εφαρμογές στην μελέτη της διαμοριακής ενέργειας αλληλεπίδρασης.



2.1. Εισαγωγή

Η πολυδιάστατη μη-γραμμική βελτιστοποίηση παίζει σημαντικό ρόλο στους *ab-initio* μοριακούς κβαντικούς υπολογισμούς καθώς απαιτεί πάρα πολύ μεγάλο υπολογιστικό χρόνο, κυρίως σε μεγάλα μοριακά συστήματα. Οι δυνάμεις Van der Waals και οι ασθενείς δεσμοί υδρογόνου που εμφανίζονται α) σε βιο-πολυμερή, β) στην διαδικασία της μοριακής αναγνώρισης και γ) στην κρυσταλλική μηχανική παίζουν καθοριστικό ρόλο στην διευθέτηση των ανόργανων και οργανικών συστημάτων^{69,70}. Η γεωμετρική και ενεργειακή περιγραφή τέτοιων μοριακών συστημάτων απαιτεί: α) την μελέτη τους με χρήση κβαντικών κυματοσυναρτήσεων μεγάλης ακρίβειας και β) την μελέτη συγκεκριμένων περιοχών της διαμοριακής υπερεπιφάνειας δυναμικής ενέργειας, ώστε να μπορούμε να εξάγουμε ποσοτικά συμπεράσματα για την φύση των ασθενών δεσμών υδρογόνου σε διμερή μοριακά συστήματα. Για την αναλυτική διερεύνηση συγκεκριμένων περιοχών της διαμοριακής υπερεπιφάνειας δυναμικής ενέργειας χρειάζονται κατάλληλα αυτοματοποιημένα υπολογιστικά εργαλεία. Ένα τέτοιο σύστημα αναπτύχθηκε και περιγράφεται στο κεφάλαιο αυτό. Αν και υπάρχουν αρκετά πακέτα μοριακής κβαντικής προσομοίωσης που παρέχουν μεθόδους γεωμετρικής αριστοποίησης, φαίνεται να χρειάζεται πιο αποδοτική και αποτελεσματική διαχείριση σε κάποιες περιοχές της διαμοριακής υπερεπιφάνειας δυναμικής ενέργειας. Η διασύνδεση του πακέτου «General Atomic and Molecular Electronic Structure System» (GAMESS-US) και του προγράμματος μη-γραμμικής βελτιστοποίησης Merlin^{71,72} έγινε με σκοπό την δημιουργία ενός γενικής χρήσης συστήματος κβαντικής χημείας το οποίο παρέχει την δική του υψηλού επιπέδου γλώσσα προγραμματισμού MCL⁷³. Το σύστημα αναπτύχθηκε με στόχο την επίλυση προβλημάτων κβαντικής χημείας τα οποία χρειάζονται ειδικές μεθόδους γεωμετρικής βελτιστοποίησης σε εξειδικευμένες περιοχές της υπερεπιφάνειας δυναμικής ενέργειας μοριακών συστημάτων και ονομάζεται MERGAM⁷⁴.

Η ανάπτυξη του συστήματος MERGAM στηρίχθηκε στην εξής προσέγγιση: το πακέτο GAMESS-US παρέχει στο MERLIN την ολική ηλεκτρονική ενέργεια ενός μοριακού συστήματος που αποτελείται από N άτομα και η διαδικασία της γεωμετρικής αριστοποίησης του συστήματος ως προς την ενέργεια ελέγχεται από το MERLIN. Το MERGAM είναι ένα πρόγραμμα που αναπτύχθηκε ως βιβλιοθήκη διασύνδεσης σε γλώσσα προγραμματισμού ANSI C, ώστε να υποστηρίξει με τον καλύτερο δυνατό τρόπο την αυτόματη επικοινωνία και ανταλλαγή δεδομένων μεταξύ των δύο πακέτων. Το κύριο χαρακτηριστικό του είναι ότι τα πακέτα GAMESS-US



και MERLIN παραμένουν ανεξάρτητα ως προς την χρήση τους και το MERGAM μπορεί να ενσωματώνεται κατά περίπτωση. Επίσης είναι άμεσα επεκτάσιμο ώστε να μπορεί να αντιμετωπίσει και άλλες ειδικές περιπτώσεις μοριακής προσομοίωσης όπως θα δούμε στο κεφάλαιο 4.

Ένα επιπλέον χαρακτηριστικό που ενσωματώθηκε στο σύστημα MERGAM είναι η δυνατότητα να υποστηρίζει την μέθοδο counterpoise (CP) των Boys and Bernardi⁷⁵. Η μέθοδος αυτή αφορά τον ακριβή υπολογισμό των διαμοριακών ενεργειών αλληλεπίδρασης διμερών μοριακών συστημάτων με μεγάλη εφαρμογή στην διερεύνηση συνήθων και ασθενών δεσμών υδρογόνου. Η μέθοδος CP εφαρμόζεται στο τέλος των ab-initio κβαντικών υπολογισμών (a-posteriori) σε διμερή μοριακά συστήματα για να εξαλείψει το λάθος που προέρχεται από την υπέρθεση των συνόλων βάσεων (Basis Set Superposition Error- BSSE) στα μονομερή που απαρτίζουν το μοριακό σύστημα. Αν το BSSE δεν ληφθεί υπ' όψη οι ενεργειακές, γεωμετρικές και λοιπές ιδιότητες του μοριακού συστήματος είναι διαστρεβλωμένες. Αν αυτή η διαστρέβλωση δεν διορθωθεί, τότε στις διαμοριακές ενέργειες ασθενών δεσμών υδρογόνου η διαδικασία της προσομοίωσης παρεκκλίνει σημαντικά από το επιθυμητό αποτέλεσμα.

Στο τέλος του παρόντος κεφαλαίου αναφέρονται δύο μελέτες σε διαφορετικά διμερή μοριακά συστήματα που έγιναν με την χρήση του συστήματος MERGAM. Τα διμερή που μελετήθηκαν είναι τα: HF-H₂O και C₆H₆-H₂O. Η επιλογή των συγκεκριμένων διμερών έγινε για τους επόμενους λόγους:

1. Για να εξερευνηθούν όλες οι δυνατότητες και χαρακτηριστικά του συστήματος MERGAM.
2. Το διμερές HF-H₂O είναι ένα καλά μελετημένο μοντέλο ώστε να εξετάσουμε την διορθωμένη ως προς το BSSE γεωμετρική αριστοποίηση στην ολική υπερεπιφάνεια δυναμικής ενέργειας.
3. Το διμερές C₆H₆-H₂O είναι ένα σύστημα που παρουσιάζει εξαιρετικό ενδιαφέρον στην μελέτη ασθενών δεσμών υδρογόνου σε βιολογικής σημασίας προσομοιώσεις μεγάλης κλίμακας. Στην περίπτωση αυτή εξετάζεται αναλυτικά η χρήση στρατηγικής βελτιστοποίησης και η διερεύνηση της διαμοριακής υπερεπιφάνειας δυναμικής ενέργειας με χρήση της υψηλού επιπέδου γλώσσας MCL που παρέχει το πακέτο MERLIN.



2.1.1. Το πακέτο GAMESS-US

Το πακέτο GAMESS-US είναι ένα από τα πιο διαδεδομένα υπολογιστικά συστήματα σε προβλήματα μοριακής κβαντικής μηχανικής. Η ανάπτυξη του ξεκίνησε από το πακέτο HONDOS καθώς και αρκετών άλλων εφαρμογών κβαντικής χημείας. Από τότε μέχρι και σήμερα συνεχώς αναπτύσσεται με νέες υπολογιστικές μεθόδους και βελτιώσεις στις υπάρχουσες εκδόσεις. Στην πράξη το GAMESS-US αποτελείται από ένα σύνολο ενοποιημένων εφαρμογών, για την μελέτη προβλημάτων κβαντικής χημείας. Πολλές από τις εφαρμογές που διαθέτει έχουν παραλληλοποιηθεί με σκοπό να έχουν μεγάλο πλεονέκτημα σε συμμετρικούς πολυεπεξεργαστές με κοινή μνήμη⁷⁶. Το πακέτο εκτελείται παράλληλα σε υπολογιστικά συστήματα με πολλούς επεξεργαστές και κοινή μνήμη (Shared Memory Processors) καθώς και σε ετερογενής υπολογιστικές αρχιτεκτονικές διασυνδεδεμένες σε τοπικό δίκτυο (Clusters).

Το GAMESS-US έχει την δυνατότητα να εκτελεί μοριακούς κβαντικούς υπολογισμούς με ένα μεγάλο πλήθος κυματοσυναρτήσεων όπως RHF, ROHF, GVB, CI και MCSCF. Η επαναληπτική διαδικασία του αυτοσυνεπούς πεδίου SCF εκτελείται με δύο τρόπους. Ο ένας τρόπος αφορά την αποθήκευση των διηλεκτρονικών ολοκληρωμάτων στον σκληρό δίσκο ώστε σε κάθε επανάληψη να επαναδιαβάζονται. Ο δεύτερος τρόπος αφορά άμεσο υπολογισμό τους σε κάθε επανάληψη, κάνοντας πλήρη χρήση της υπολογιστικής μνήμης και της επεξεργαστικής ισχύος.

Για τον υπολογισμό της δυναμικής ηλεκτρονικής ενέργειας συσχέτισης στα μοριακά συστήματα χρησιμοποιεί τις εξής μεθόδους: την αλληλεπίδραση ηλεκτρονικών σχηματισμών (Configuration Interaction), θεωρία διαταραχών δεύτερης τάξης, και προσέγγιση couple-cluster. Επίσης παρέχει υπολογισμούς με την θεωρία συναρτησιακών ηλεκτρονικής πυκνότητα (Density Functional Theory – DFT). Το πακέτο διαθέτει αναλυτικούς παραγώγους της ολικής ηλεκτρονικής ενέργειας ως προς την γεωμετρία ώστε να παρέχει γεωμετρική αριστοποίηση και εύρεση μεταπτώσιμων καταστάσεων μοριακών συστημάτων. Ο υπολογισμός των δεύτερων παραγώγων ως προς την ενέργεια χρησιμοποιείται για την προσομοίωση των δονητικών συχνοτήτων και φυσικοχημικών ιδιοτήτων ενός μοριακού συστήματος. Διάφορες μοριακές ιδιότητες όπως διπολική ροπή και πολωσιμότητες μπορούν να υπολογιστούν. Το πακέτο GAMESS-US έχει ενσωματωμένο ένα μεγάλο πλήθος συνόλων βάσεων που περιγράφουν τα ατομικά τροχιακά των ατόμων.

Το GAMESS-US διανέμεται δωρεάν στην ακαδημαϊκή κοινότητα και υποστηρίζει



όλα τα γνωστά λειτουργικά συστήματα. Το πακέτο δίνεται σε μορφή ανοιχτού κώδικα καθώς και σε εκτελέσιμη μορφή. Συνοδεύεται από αναλυτικό εγχειρίδιο και αρκετά παραδείγματα που αξιοποιούν τις πιο σημαντικές δυνατότητές του.

2.1.2. Το περιβάλλον μη-γραμμικής βελτιστοποίησης MERLIN

Το MERLIN⁷¹ είναι ένα περιβάλλον πολυδιάστατης μη-γραμμικής βελτιστοποίησης γενικής χρήσης. Αναπτύχθηκε για να αντιμετωπίζει πολυδιάστατα μη-γραμμικά προβλήματα της μορφής:

$$\text{Εύρεση ενός τοπικού ελάχιστου της συνάρτησης } f(x), x \in R^N, x = [x_1, x_2, \dots, x_N]^T,$$

$$\text{υπό τους περιορισμούς: } x_i \in [a_i, b_i] \text{ για } i = 1, 2, \dots, N.$$

Η συνάρτηση $f(x)$ καλείται αντικειμενική συνάρτηση και οι παράμετροι x_i αποτελούν ένα σημείο στον πολυδιάστατο χώρο.

Το περιβάλλον MERLIN παρέχει μια σειρά από αλγόριθμους βελτιστοποίησης τους οποίους μπορεί να επιλέξει ο χρήστης και να εφαρμόσει στο πρόβλημα που επιθυμεί να επιλύσει. Οι αλγόριθμοι αυτοί χρησιμοποιούν κατά περίπτωση αριθμητικές ή αναλυτικές πρώτες παραγώγους κατά την διάρκεια της βελτιστοποίησης. Στις εφαρμογές του κεφαλαίου αυτού με το σύστημα MERGAM χρησιμοποιήθηκε η μέθοδος Tolmin^{77,78}.

Το MERLIN παρέχει το δικό του «φιλικό» περιβάλλον διεπαφής με τον χρήστη και προσφέρει δυνατότητα επιλογής αλγορίθμων βελτιστοποίησης ανάλογα με το μελετούμενο σύστημα, μέσω εντολών που δίνει ο χρήστης. Ολόκληρο το πακέτο MERLIN είναι γραμμένο ως υπορουτίνα σε γλώσσα προγραμματισμού Fortran 77. Ο χρήστης καλείται να προγραμματίσει το κυρίως πρόγραμμα που στην ουσία είναι η αντικειμενική συνάρτησης που θα βελτιστοποιηθεί. Προαιρετικά και εφόσον χρειάζεται έχει την δυνατότητα να προγραμματίσει τις πρώτες ή δεύτερες παραγώγους της αντικειμενικής συνάρτησης.

Ένα σημαντικό χαρακτηριστικό του πακέτου MERLIN είναι οι δυνατότητες επέκτασης που παρέχει. Με αυτό τον τρόπο νέες δυνατότητες μπορούν να εισαχθούν στο πακέτο ως επεκτάσεις "plugins", κάνοντας χρήση της βιβλιοθήκης εσωτερικών συναρτήσεων και ρουτινών που παρέχει. Οι ρουτίνες αυτές ("glue routines") είναι αναπόσπαστο κομμάτι του MERLIN και παρέχουν πρόσβαση στις εσωτερικές δομές του. Η λειτουργικότητα του πακέτου ολοκληρώνεται με τις δυνατότητες που προσφέρει στην τεχνικές βελτιστοποίησης διαμέσου της δικής του γλώσσας



προγραμματισμού υψηλού επιπέδου “Merlin Control Language”⁷³. Η MCL έχει πλούσιο συντακτικό και περιλαμβάνει όλα τα χαρακτηριστικά μιας διαδικαστικής γλώσσας προγραμματισμού όπως: ανάθεση, επανάληψη, έλεγχο, διαχείριση αρχείων, είσοδο/έξοδο και υποπρογράμματα. Η MCL είναι φτιαγμένη ειδικά ώστε να αλληλεπιδρά άμεσα με τις εσωτερικές δομές, λειτουργίες και επεκτάσεις του MERLIN ώστε να καθιστά εφικτές διαδικασίες απρόσκοπτης βελτιστοποίησης που χρειάζονται για να επιτευχθεί ένας στρατηγικός στόχος. Οι αναλυτικές τεχνικές λεπτομέρειες για την επέκταση καθώς και η πλήρης περιγραφή των εντολών του πακέτου MERLIN/MCL αναπτύσσονται στις δημοσιεύσεις και τεχνικές οδηγίες του συστήματος^{71,72,73,79}.

2.2. Διασύνδεση και αναλυτική περιγραφή του συστήματος MERCAM

Όλες οι κβαντικές μοριακές προσομοιώσεις απαιτούν τις 3N καρτεσιανές συντεταγμένες που περιγράφουν τις θέσεις των N ατόμων που απαρτίζουν το σύστημα. Ο υπολογισμός της ολικής ηλεκτρονικής ενέργειας ενός μοριακού συστήματος με το πακέτο GAMESS-US γίνεται: α) ορίζοντας την γεωμετρία ενός μοριακού συστήματος, β) επιλέγοντας το επίπεδο θεωρίας των κβαντικών υπολογισμών και γ) επιλέγοντας το σύνολο βάσεων για τα ατομικά τροχιακά. Στο σύστημα MERCAM που αναπτύχθηκε, η διαχείριση των γεωμετρικών χαρακτηριστικών του μοριακού συστήματος γίνεται με το πακέτο MERLIN/MCL και ο υπολογισμών των διαφόρων ενεργειακών ιδιοτήτων γίνεται με το πακέτο GAMESS-US. Η προσέγγιση που ακολουθήθηκε είναι η εξής: Το MERLIN παρέχει στο GAMESS-US την μοριακή γεωμετρία, το GAMESS-US υπολογίζει τις ενεργειακές ιδιότητες και τις επιστρέφει στο MERLIN.

Ο χρήστης δημιουργεί ένα αρχείο εισόδου, προδιαγεγραμμένο για υπολογισμό με το σύστημα GAMESS-US και η γεωμετρική αριστοποίηση ελέγχεται από το περιβάλλον MERLIN/MCL. Στην πραγματικότητα η αντικειμενική συνάρτηση καθώς και οι πρώτες – δεύτερες αναλυτικές παράγωγοι γίνονται με κλήση στο πακέτο GAMESS-US.

Στην συνέχεια θα αναλύσουμε την αρχιτεκτονική, λειτουργικές προδιαγραφές και αναλυτικές τεχνικές λεπτομέρειες της ανάπτυξης του συστήματος MERCAM. Για λόγους απλοποίησης στις επόμενες ενότητες θα αναφέρουμε τις συναρτήσεις και υπορουτίνες της γλώσσας προγραμματισμού Fortran 77 ως ρουτίνες.



2.2.1. Η αρχιτεκτονική και η επικοινωνία του συστήματος

Η ανάπτυξη ενός σταθερού και διαλειτουργικού συστήματος επικοινωνίας μεταξύ των πακέτων GAMESS-US και MERLIN/MCL έγινε με την ανάπτυξη κώδικα σε γλώσσα προγραμματισμού ANSI-C. Ο κώδικας αναπτύχθηκε ως βιβλιοθήκη διασύνδεσης (library) μεταξύ των δύο πακέτων και συνδέεται με τον εκτελέσιμο κώδικα του MERLIN μέσω των προγραμματιζόμενων ρουτινών που παρέχει στον χρήστη. Η διαδικασία αυτή διατηρεί την ανεξαρτησία των δύο πακέτων και δεν εμπλέκει τον τελικό χρήστη με τεχνικές λεπτομέρειες.

Ο πηγαίος κώδικας της βιβλιοθήκης που αναπτύχθηκε και είναι ο βασικό μηχανισμός λειτουργίας του συστήματος, βρίσκεται στο αρχείο "gms_ari.c" που διανέμεται με το σύστημα MERCAM.

Οι συναρτήσεις που αναπτύχθηκαν έχουν ως κύριο στόχο:

- a. Την διαχείριση των αρχείων δεδομένων του πακέτου GAMESS-US και
- b. Την απρόσκοπτη δι-επικοινωνία μεταξύ των δύο πακέτων μέσω κλήσεων σε συναρτήσεις του λειτουργικού συστήματος.

Η αναλυτική αρχιτεκτονική του συστήματος καθώς και η αλληλουχία κλήσεων φαίνεται στην εικόνα 2.1.

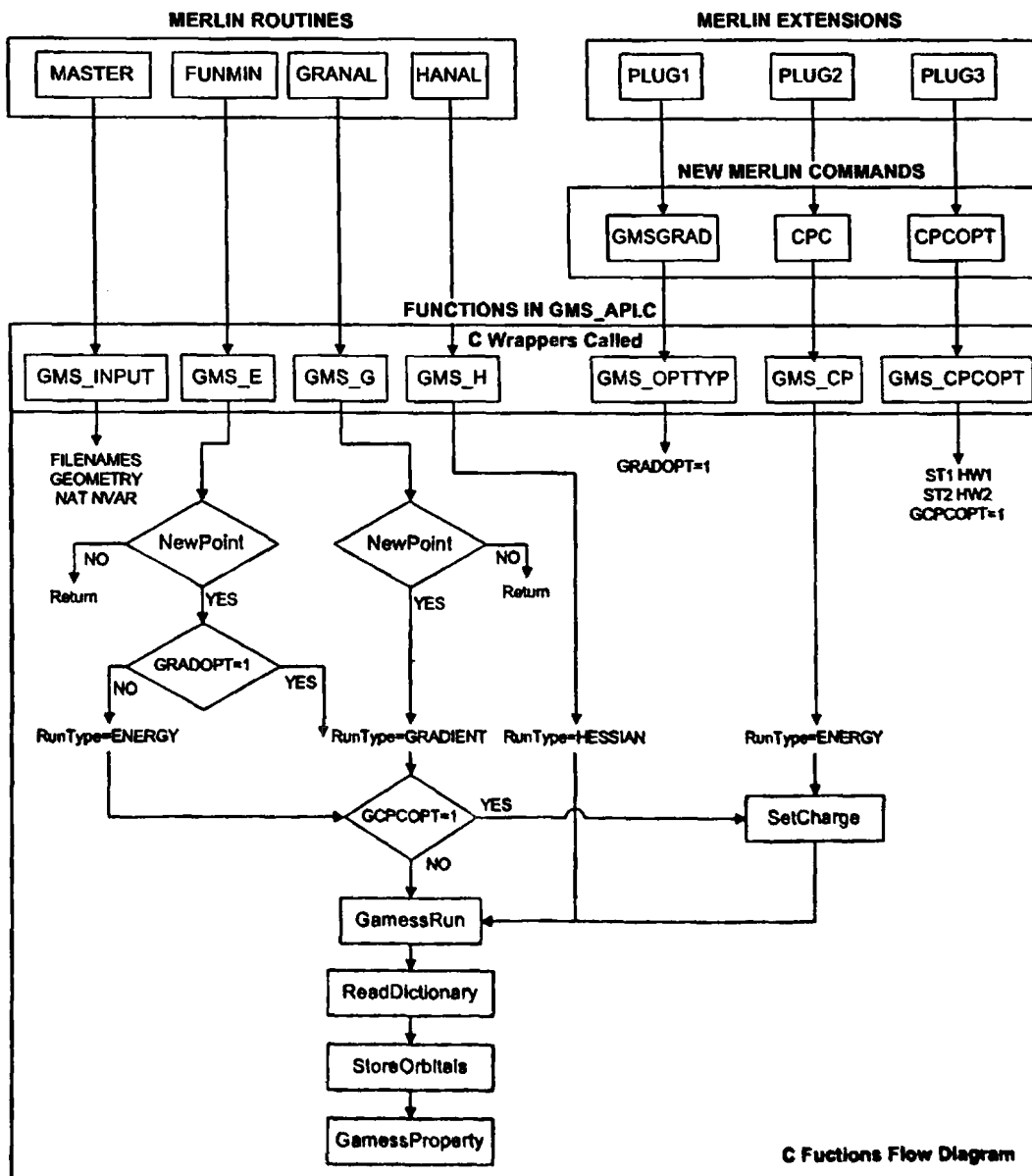
Στην αρχιτεκτονική του συστήματος MERCAM στην εικόνα 2.1 είναι φανερό ότι το κύριο σύστημα διαχείρισης είναι το πακέτο MERLIN, το οποίο παρέχει ένα πολύ αξιόπιστο περιβάλλον διεπαφής στον χρήστη. Το πακέτο MERLIN/MCL επεκτάθηκε με νέες ρουτίνες και νέες εντολές. Η επικοινωνία και ανταλλαγή δεδομένων μεταξύ των δύο πακέτων γίνεται με την χρήση ειδικών συναρτήσεων σε C, οι οποίες ονομάζονται C-wrappers. Ο κώδικας που αναπτύχθηκε για την διασύνδεση χωρίστηκε σε δύο λογικά υποσύνολα συναρτήσεων: α) C-wrappers και β) C-primitives.

Οι C-wrappers είναι συναρτήσεις που αναδιπλώνουν ή καλύτερα «κρύβουν» τον κώδικα που αφορά συγκεκριμένο μεταγλωττιστή (compiler) και λειτουργικό σύστημα. Ο λόγος για τον οποίο αναπτύχθηκαν είναι καθαρά τεχνικός και παρέχει διαλειτουργικότητα σε λειτουργικά συστήματα με μεταγλωττιστές C και Fortran 77 διαφορετικών κατασκευαστών. Με αυτό τον τρόπο το σύστημα επιτυχώς εγκαθίσταται και εκτελείται σε διαφορετικές αρχιτεκτονικές υπολογιστικών συστημάτων προσομοιώσεων ακόμα και αν πληρούν εξειδικευμένες προδιαγραφές. Για το λόγο αυτό το MERLIN επικοινωνεί μόνο μέσω των C-wrappers με το GAMESS-US παρέχοντας έμπιστο και ασφαλή τρόπο διεπικοινωνίας και ανταλλαγής δεδομένων. Οι C-wrappers λειτουργούν ως ενδιάμεσες συναρτήσεις επικοινωνίας και



είναι αναγκαίες όταν απαιτείται επικοινωνία προγραμμάτων που είναι γραμμένα σε διαφορετικές γλώσσες προγραμματισμού. Όλες οι συναρτήσεις C-wrappers παρέχουν την ασφαλή διεπαφή μεταξύ των δυο πακέτων και βρίσκονται στο πηγαίο αρχείο "mer_ext.h" υποστηρίζοντας όλους τους γνωστούς μεταγλωττιστές της γλώσσας προγραμματισμού C.

Εικόνα 2.1



Η αρχιτεκτονική του MERGAM

Το υποσύνολο συναρτήσεων C-primitives, καλείται μόνο από συναρτήσεις C-wrappers. Αφορά την εκτέλεση λειτουργιών που σχετίζονται με την εκτέλεση του GAMESS-US και την πρόσβαση στις εξωτερικές δομές αποθήκευσης πληροφοριών



που αφορούν τους κβαντικούς μοριακούς υπολογισμούς. Η πιο σημαντική συνάρτηση του συνόλου των C-primitives είναι η συνάρτηση GamesProperty στην οποία «καταλήγει» η αρχιτεκτονική του συστήματος MERCAM, ενώ στην συνέχεια παρέχει τα αποτελέσματα των κβαντικών υπολογισμών στους C-wrappers ώστε με την σειρά τους να «περάσουν» στο MERLIN.

Η αναλυτική περιγραφή των ρουτινών C-primitives και C-wrappers βρίσκεται στα παραρτήματα Α και Β αντίστοιχα. Οι τρεις τελευταίες συναρτήσεις C-wrappers του παραρτήματος Β υλοποιήθηκαν ως νέες του εντολές του περιβάλλοντος MERLIN/MCL χρησιμοποιώντας την τεχνολογία των επεκτάσεων plug-in που διαθέτει το πακέτο MERLIN. Όπως φαίνεται και από την εικόνα 2.1 της αρχιτεκτονικής οι C-wrappers βρίσκονται σε υψηλότερο λογικό επίπεδο από τις C-primitives. Σε αυτή την λογική αλληλουχία στηρίχθηκε η αυτοματοποίηση και η διαδικασία της διασύνδεσης, παρέχοντας και τα χαρακτηριστικά της εύκολης και σταθερής προσαρμογής - επέκτασης του συστήματος.

Το διασυνδεδεμένο σύστημα MERCAM λειτουργεί κάνοντας χρήση προσωρινών αρχείων και μεταβλητών περιβάλλοντος σε λειτουργικό σύστημα UNIX. Τα αρχεία χρησιμοποιούνται για την αποθήκευση δεδομένων και αποτελούν τον πιο αποτελεσματικό τρόπο ώστε τα δεδομένα να διαμοιράζονται μεταξύ διαφορετικών υπολογιστικών εφαρμογών. Στις περισσότερες και συνήθεις υπολογιστικές αρχιτεκτονικές ένα πρόγραμμα εκτελείται αποκλειστικά από τον επεξεργαστή και συνήθως αυτό είναι επιθυμητό ώστε να υπάρχει η μέγιστη δυνατή απόδοση. Κατά την διάρκεια εκτέλεσης ενός προγράμματος η μνήμη του υπολογιστή καταλαμβάνεται από δεδομένα τα οποία αποθηκεύονται προσωρινά, όσο διαρκεί και η εκτέλεση του προγράμματος, με μοναδικό πλεονέκτημα την ταχύτητα πρόσβασης. Τα τελικά δεδομένα της εκτέλεσης ενός προγράμματος αποθηκεύονται σε αρχεία τα οποία αποτελούν λογικές οντότητες που διατηρούνται σε αποθηκευτικά μέσα όπως ο σκληρός δίσκος. Τα αρχεία αυτά είναι κοινά για το ίδιο λειτουργικό σύστημα και μπορούν όλες οι εφαρμογές να έχουν πρόσβαση σε αυτά. Την διαδικασία αυτή ακολουθεί και το σύστημα MERCAM. Ο κώδικας που αναπτύχθηκε αποτελεί επέκταση του MERLIN και εκτελείται εφόσον εκτελείται το MERLIN και μόνο. Ο κώδικας του MERCAM γράφει σε αρχεία όλες τις απαραίτητες πληροφορίες που χρειάζεται το GAMESS-US για να εκτελέσει ένα υπολογισμό προσομοίωσης. Όταν ο υπολογισμός του GAMESS-US ολοκληρωθεί τότε όλα τα δεδομένα αποθηκεύονται σε αρχεία. Το MERCAM διαβάζει αυτά τα αρχεία και τα δεδομένα είναι διαθέσιμα



για επεξεργασία στο MERLIN.

Για να επιτευχθεί η διαδικασία που περιγράψαμε πρέπει το GAMESS-US να γνωρίζει ποια αρχεία θα χρησιμοποιήσει και το MERGAM ποια αρχεία αποτελεσμάτων θα διαβάσει. Σε αυτό το σημείο γίνεται χρήση των μεταβλητών περιβάλλοντος (environment variables) που διαθέτει κάθε λειτουργικό σύστημα. Ένα περιγραφικό αρχείο κελύφους (shell script) χρησιμοποιείται στο MERGAM και η εκκίνηση του συστήματος γίνεται με αυτό. Το αρχείο αυτό περιέχει σε μεταβλητές περιβάλλοντος τα ονόματα και την θέση των αρχείων που θα χρησιμοποιηθούν για την επικοινωνία. Κάθε πρόγραμμα που εκτελείται σε ένα λειτουργικό σύστημα έχει πρόσβαση στις μεταβλητές περιβάλλοντος τις οποίες διαβάζει ώστε να πάρει πληροφορίες για εξωτερικά δεδομένα που σχετίζονται με την εκτέλεσή του.

2.2.2. Ο κυρίως κώδικας του συστήματος

Η διασύνδεση μεταξύ των πακέτων GAMESS-US και MERLIN/MCL επιτεύχθηκε μέσω των ρουτινών: MASTER (κυρίως πρόγραμμα), FUNMIN, GRANAL και HANAL, που διαθέτει το MERLIN στον χρήστη για προγραμματισμό. Με αυτό τον τρόπο η μελέτη ενός προβλήματος ανάγεται στον προγραμματισμό των εξής:

- Του κυρίως προγράμματος,
- Της αντικειμενικής συνάρτησης,
- Των αναλυτικών πρώτων και
- Των αναλυτικών δεύτερων παραγώγων αντίστοιχα.

Οποιαδήποτε εφαρμογή με το MERLIN απαιτεί τον προγραμματισμό αυτών των συναρτήσεων από τον χρήστη. Η ανάπτυξη του MERGAM ξεκίνησε και στηρίχθηκε στον προγραμματισμό αυτών των ρουτινών. Το MERLIN χρησιμοποιεί μόνο αυτές τις ρουτίνες για να επικοινωνήσει με το GAMESS-US και η ανταλλαγή δεδομένων γίνεται μέσω των τυπικών παραμέτρων των ρουτινών.

Ο πηγαίος κώδικας των ρουτινών βρίσκεται στο αρχείο "mergam.d". Το σύστημα MERGAM ξεκινάει την λειτουργία του από την ρουτίνα MASTER και οποιαδήποτε κλήση στο GAMESS-US γίνεται ως επί το πλείστον από τις υπορουτίνες FUNMIN, GRANAL και HANAL, όπως φαίνεται και στην εικόνα 2.1. Μέσα στο αρχείο "mergam.d" υπάρχουν επίσης μακρο-μεταβλητές εγκατάστασης \$(TYPE) και \$(MAXW) οι οποίες υποκαθίστανται κατά την εγκατάσταση του MERLIN με τον τύπο ακρίβειας των πραγματικών αριθμών που χρησιμοποιεί το σύστημα και τον



μέγιστο αριθμό μεταβλητών αντίστοιχα.

Στην συνέχεια παρουσιάζονται αναλυτικά οι προγραμματιζόμενες ρουτίνες του MERLIN/MCL και περιγράφεται η λειτουργία τους:

- **PROGRAM MASTER:** Είναι το σημείο της αρχής εκτέλεσης του προγράμματος. Η C-wrapper συνάρτηση GMS_INPUT καλείται για να αρχικοποιήσει τις μεταβλητές επικοινωνίας του συστήματος. Η συνάρτηση GMS_NPUT διαβάζει το αρχείο εισόδου του πακέτου GAMESS-US και το αποθηκεύει στην μνήμη μαζί με τον συνολικό αριθμό των καρτεσιανών ατομικών συντεταγμένων.
- **FUNCTION FUNMIN(x, n):** Η συνάρτηση αυτή καλείται αυτόματα από το MERLIN όποτε χρειάζεται να γίνει υπολογισμός της αντικειμενικής συνάρτησης. Στην περίπτωση μας είναι η ολική ηλεκτρονική ενέργεια που παρέχει ένας υπολογισμός με το GAUSSIAN-US. Οι παράμετροι κατά τιμή με τις οποίες τροφοδοτείται η συνάρτηση είναι οι ατομικές καρτεσιανές συντεταγμένες πλήθους N. Η συνάρτηση FUNMIN καλεί την C-wrapper συνάρτηση GMS_E με τις ίδιες παραμέτρους. Η GMS_E εκτελεί το GAMESS-US στο σημείο X(N) και επιστρέφει την ολική ηλεκτρονική ενέργεια στην FUNMIN και κατ' επέκταση στο MERLIN.
- **SUBROUTINE GRANAL(N, X, GRAD):** Η ρουτίνα GRANAL του MERLIN καλείται για τον υπολογισμό των πρώτων αναλυτικών παραγώγων. Οι παράμετροι κατά τιμή X και N είναι οι καρτεσιανές ατομικές συντεταγμένες ενώ η μεταβλητή GRAD αναπαριστά κατά διεύθυνση έναν μονοδιάστατο πίνακα διάστασης N που περιέχει τις πρώτες παραγώγους της αντικειμενικής συνάρτησης. Η υπορουτίνα GRANAL καλεί την C-wrapper συνάρτηση GMS_G η οποία εκτελεί έναν υπολογισμό GRADIENT στο GAMESS-US και επιστρέφει την ενέργεια και τις πρώτες παραγώγους της ενέργειας ως προς την γεωμετρία. Η υπορουτίνα GRANAL καλείται από το MERLIN εφόσον έχει εκτελεστεί η εντολή του MERLIN "anal". Αν δεν έχει εκτελεστεί η εντολή "anal" οι πρώτες παράγωγοι υπολογίζονται αριθμητικά με χρήση της ρουτίνας FUNMIN.
- **SUBROUTINE HANAL(H, LD, N, X):** Η ρουτίνα αυτή καλείται από το MERLIN όποτε χρειάζονται οι δεύτερες αναλυτικές παράγωγοι ή καλύτερα ο εσσιανός πίνακας. Το MERLIN περνάει κατά τιμή τις ατομικές συντεταγμένες X(N) καθώς και την μέγιστη διάσταση LD του πίνακα. Ο δυοδιάστατος N*N πίνακας που αντιπροσωπεύεται από την μεταβλητή H επιστρέφει της δεύτερες αναλυτικές παραγώγους στο MERLIN. Η υπορουτίνα HANAL καλεί την C-wrapper



συνάρτηση GMS_H η οποία με την σειρά της εκτελεί έναν υπολογισμό τύπου HESSIAN στο GAMESS-US για τον υπολογισμό των δεύτερων παραγώγων της ολικής ηλεκτρονικής ενέργειας του μοριακού συστήματος ως προς την γεωμετρία. Το MERLIN, όπως προαναφέρθηκε, χρησιμοποιεί τις παραπάνω ρουτίνες για να κάνει όλους τους υπολογισμούς που σχετίζονται με την διαδικασία της βελτιστοποίησης. Στην περίπτωση του συστήματος MERGAM, το MERLIN πρέπει να «δώσει» δεδομένα στο GAMESS-US και να πάρει αποτελέσματα. Ο τρόπος με τον οποίο υλοποιείται αυτή η διαδικασία αναπτύχθηκε στο MERGAM και είναι η εξής:

- Το MERLIN παρέχει ένα σημείο X(N) της επιφάνειας δυναμικής ενέργειας στο οποίο πρέπει να γίνει ο υπολογισμός της ολικής ηλεκτρονικής ενέργειας με το GAMESS-US.
- Το σημείο αυτό, μέσω των C-wrappers, γράφεται σε ένα αρχείο εισόδου (input file) του GAMESS-US. Η δημιουργία του αρχείου εισόδου γίνεται με κλήση συνάρτησης τύπου C-primitive.
- Το GAMESS-US εκτελείται με αυτό το αρχείο εισόδου.
- Ο υπολογισμός ολοκληρώνεται.
- Η κατάλληλη συνάρτηση C-primitive καλείται για να διαβάσει το δυαδικό αρχείο αποτελεσμάτων του GAMESS-US.
- Τα αποτελέσματα μεταφέρονται μέσω των C-wrappers σε μεταβλητές του MERLIN.

Ο κώδικας του MERGAM αποτελείται από τους κατάλληλους C-wrappers οι οποίοι εκτελούν πολλαπλούς υπολογισμούς στο GAMESS-US, αποθηκεύουν τα αποτελέσματα σε μεταβλητές, κάνουν κατάλληλες πράξεις που αντιπροσωπεύουν μία φυσική ποσότητα και μεταφέρουν την τελική τιμή στο MERLIN. Με αυτό τον τρόπο μπορούν να αυτοματοποιηθούν πολύπλοκες υπολογιστικές διαδικασίες όπως θα δούμε και στην συνέχεια.



2.3. Εφαρμογή του περιβάλλοντος MERCAM σε εξειδικευμένα προβλήματα κβαντικής μοριακής μηχανικής

Το σύστημα MERCAM μπορεί να εφαρμοστεί σε ένα σύνολο από προβλήματα υπολογιστικής προσομοίωσης που αφορούν την μοριακή κβαντική μηχανική. Τα προβλήματα αυτά καλύπτουν ως επί το πλείστον την γεωμετρική βελτιστοποίηση μοριακών συστημάτων, την διερεύνηση του διαμορφωσιακού χώρου και την εξερεύνηση της επιφάνειας δυναμικής ενέργειας. Το κύριο και πιο σημαντικό χαρακτηριστικό του MERCAM είναι η δυνατότητα που παρέχει να προγραμματιστούν υπολογιστικές διαδικασίες σε ένα ή περισσότερα σημεία της επιφάνειας δυναμικής ενέργειας. Για τον λόγο αυτό στην παρούσα διατριβή επιλέχθηκε ένα αρκετά πολύπλοκο πρόβλημα το οποίο επιλύθηκε αυτοματοποιημένα. Το πρόβλημα αφορά την μελέτη της διαμοριακής επιφάνειας δυναμικής ενέργειας ασθενώς αλληλεπιδρώντων μοριακών συστημάτων, ενός σημαντικού υποσυνόλου της συνολικής μοριακής επιφάνειας δυναμικής ενέργειας.

Γενικότερα, το MERCAM μπορεί να υποστηρίξει τα επόμενα γενικά και ειδικά προβλήματα μοριακής κβαντικής μηχανικής:

1. Γεωμετρική βελτιστοποίηση με ποικιλία μεθόδων.
2. Γεωμετρική βελτιστοποίηση θέτοντας γεωμετρικούς περιορισμούς στο μοριακό σύστημα.
3. Βελτιστοποίηση της διαμοριακής επιφάνειας δυναμικής ενέργειας.
4. Υπολογισμός φυσικών ποσοτήτων όπως το BSSE.
5. Υλοποίηση σάρωσης με περιορισμένους ενδομοριακούς ή διαμοριακούς βαθμούς ελευθερίας.
6. Αυτοματοποιημένη μελέτη μιας περιοχής της επιφάνειας διαμοριακής ενέργειας όπου ένα μοριακό σύστημα υπόκειται σε κίνηση που καθορίζεται από μαθηματική εξίσωση.
7. Αυτοματοποιημένη στρατηγική μελέτη με συνδυασμό όλων των παραπάνω.

Από το πλήθος των πακέτων μοριακής κβαντικής μηχανικής που διατίθενται για επιστημονικούς και ερευνητικούς σκοπούς τα προβλήματα 6 και 7 αποτελούν μοναδικά χαρακτηριστικά του συστήματος MERCAM.

2.3.1. Η διαμοριακή ενέργεια αλληλεπίδρασης

Οι διαμοριακές αλληλεπιδράσεις αναφέρονται στις φυσικές δυνάμεις που ασκούνται μεταξύ μορίων με σκοπό την δημιουργία μακροσκοπικών μοριακών συστημάτων



(υγρά, διαλύματα, μοριακά στερεά, κλπ) και μικροσκοπικών μοριακών κατασκευών (διμερή, τριμερή, μοριακές συστοιχίες κλπ)⁸⁰. Σε αντίθεση με τις ενδομοριακές δυνάμεις, οι οποίες συγκρατούν τα άτομα μεταξύ τους και δημιουργούν ισχυρούς δεσμούς, οι διαμοριακές δυνάμεις είναι πολύ ασθενέστερες από τις ενδομοριακές⁸¹. Οι διαμοριακές αλληλεπιδράσεις οφείλονται κυρίως σε ισχυρούς δεσμούς υδρογόνου, ασθενείς δεσμούς υδρογόνου και δεσμούς δυνάμεων van der Waals. Στον επόμενο πίνακα φαίνεται η τάξη μεγέθους των διαφόρων δυνάμεων που εμφανίζονται στους δεσμούς των μοριακών συστημάτων.

Πίνακας

Τύπος Δεσμού	Τάξης Μεγέθους
Ιοντικός	1000
Δεσμός Υδρογόνου	100
Δίπολο-Δίπολο	10
van der Waals (London Forces)	1

Η δομική, ενεργειακή και δυναμική συμπεριφορά των μοριακών συμπλόκων εξαρτάται από το διαμοριακό δυναμικό το οποίο περιγράφει την ενέργεια και την κατεύθυνση των διαμοριακών δεσμών. Ο υπολογισμός του διαμοριακού δυναμικού είναι το πιο σημαντικό στάδιο ώστε να κατανοήσουμε αυτές τις ιδιότητες.

Η ολική ενέργεια E μιας συλλογής μορίων που αλληλεπιδρούν μεταξύ τους είναι:

$$E = \sum_i E_i + \sum_{i>j} V_{ij} + \sum_{i>j>k} V_{ijk} + \dots \quad (2.1)$$

όπου: E_i είναι η ολική ενέργεια του μορίου i , V_{ij} είναι η ενέργεια της αλληλεπίδρασης μεταξύ μόνο των μορίων i και j , V_{ijk} είναι η ενέργεια της αλληλεπίδρασης μεταξύ μόνο των μορίων i, j, k κ.ο.κ.

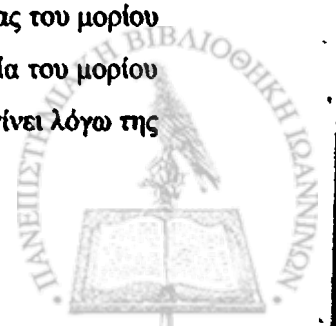
Η διαμοριακή ενέργεια αλληλεπίδρασης ορίζεται ως:

$$V = E - \sum_i E_i = \sum_{i>j} V_{ij} + \sum_{i>j>k} V_{ijk} + \dots \quad (2.2)$$

Στην περίπτωση που έχουμε τρία μόρια A, B και C η εξίσωση (2.2) γίνεται:

$$V = V_{AB} + V_{BC} + V_{CA} + V_{ABC} \quad (2.3)$$

όπου V_{AB} είναι η ενέργεια αλληλεπίδρασης των μορίων A και B απουσίας του μορίου C . Η ενέργεια μεταξύ των μορίων A και B θα ήταν διαφορετική παρουσία του μορίου C . Ο μη-προσθετικός όρος V_{ABC} εκφράζει την διόρθωση που πρέπει να γίνει λόγω της



απουσίας των C, A και B μορίων από τους υπολογισμούς των ενεργειών αλληλεπίδρασης των όρων V_{AB} , V_{BC} και V_{CA} της εξίσωσης (2.3) αντίστοιχα⁸⁰.

Οι διαμοριακές δυνάμεις που ασκούνται μεταξύ δύο ή περισσότερων μορίων είναι συνάρτηση της διαμοριακής απόστασης R μεταξύ των μορίων. Ανάλογα με την διαμοριακή απόσταση R οι δυνάμεις ανήκουν είτε στην κατηγορία μεγάλης εμβέλειας (long-range) είτε στην κατηγορία μικρής εμβέλειας (short-range). Τρεις τύποι διαμοριακών δυνάμεων ανήκουν στην κατηγορία long-range: α) οι ηλεκτροστατικές (electrostatic), β) οι επαγωγικές (induction) και γ) οι διασποράς (dispersion, van der Waals, London). Στην κατηγορία short-range ανήκουν δύο τύποι διαμοριακών δυνάμεων: ε) οι ανταλλαγής (exchange) και στ) οι απωστικές (repulsion). Οι δύο τύποι των short-range ενοποιούνται σε έναν κοινό τύπο exchange-repulsion. Η ολική διαμοριακή ενέργεια αλληλεπίδρασης αποτελείται και από τους πέντε προηγούμενους τύπους.

Ο υπολογισμός της διαμοριακής ενέργειας αλληλεπίδρασης εξαρτάται από τον σχετικό προσανατολισμό των αλληλεπιδρώντων μορίων. Στην περίπτωση μη-γραμμικών μορίων η διαμοριακή ενέργεια αλληλεπίδρασης ορίζεται ως δυναμική επιφάνεια διαμοριακής ενέργειας πολλών μεταβλητών. Υπάρχουν αρκετά σημεία ελάχιστης ενέργειας στην επιφάνεια αυτή τα οποία είναι τοπικά ελάχιστα και ένα ολικό ελάχιστο. Η διαμοριακή ενέργεια αλληλεπίδρασης εξαρτάται από την απόσταση μεταξύ των κέντρων μάζας των μορίων καθώς και από τις έξι γωνίες Euler που απαιτούνται για να καθορίσουν τον σχετικό προσανατολισμό του συστήματος αναφοράς των μορίων ως προς το ολικό σύστημα συντεταγμένων.

2.3.1.1. Θεωρητικές μέθοδοι υπολογισμού της διαμοριακής ενέργειας αλληλεπίδρασης

Οι διαμοριακές δυνάμεις είναι γενικά πολύ ασθενείς δυνάμεις σε σχέση με τους ισχυρούς δεσμούς που συγκρατούν τα άτομα μεταξύ τους. Ο ακριβής υπολογισμός της διαμοριακής ενέργειας αλληλεπίδρασης απαιτεί ως επί το πλείστον τη χρήση θεωρίας διαταραχών (perturbation theory). Αν τα μόρια A και B ενός διμερούς συστήματος είναι αρκετά μακριά τότε η θεωρία υπολογισμού της διαμοριακής ενέργειας γίνεται σχετικά απλή, επειδή η αλληλοεπικάλυψη μεταξύ των κυματοσυναρτήσεων των μορίων αμελείται^{82,83}.

Στην περίπτωση που χρησιμοποιείται κβαντική θεωρία διαταραχών για τον υπολογισμό των διαμοριακών ενεργειών οι ηλεκτροστατικές και οι δυνάμεις



ανταλλαγής-άπωσης εμφανίζονται στους όρους πρώτης τάξης (first order), ενώ οι δυνάμεις επαγωγής και διασποράς εμφανίζονται στους όρους δεύτερης τάξης(second order)⁸⁰. Η διαμοριακή γεωμετρία των ασθενών van der Waals συμπλόκων εξαρτάται περισσότερο από τις ηλεκτροστατικές δυνάμεις οι οποίες είναι πιο ισχυρές ακόμα και αν δεν έχουν την επικρατέστερη συνεισφορά στην ολική ενέργεια αλληλεπίδρασης. Οι δεσμοί van der Waals καθώς και οι ασθενείς δεσμοί υδρογόνου εξαρτώνται από την ενέργεια ηλεκτρονικής συσχέτισης σε τόσο μεγάλο βαθμό ώστε η θεωρία Hartree-Fock αποτυγχάνει να τους περιγράψει⁸¹. Το ίδιο πρόβλημα υπάρχει και με την θεωρία DFT η οποία δεν μπορεί να περιγράψει ηλεκτρονική συσχέτιση τύπου 'non-local'^{84,85}. Η θεωρία διαταραχών Moller-Plesset δεύτερης τάξης αποδεικνύεται να είναι μια πολύ καλή μέθοδος με σχετικά μικρό υπολογιστικό κόστος για τον υπολογισμό ασθενών αλληλεπιδράσεων διαμοριακών συμπλόκων. Η θεωρία CCSD(T) αποδίδει καλύτερα και έχει μεγάλο κόστος ακόμα και για μεσαίας τάξης προβλήματα.

Στην πράξη υπάρχουν δύο υπολογιστικές μέθοδοι για την μέτρηση της διαμοριακής ενέργειας αλληλεπίδρασης. Η πρώτη μέθοδος ονομάζεται θεωρία διαταραχών προσαρμοσμένη στην συμμετρία (SAPT – Symmetry Adapted Perturbation Theory)^{86,87}. Η δεύτερη μέθοδος η οποία χρησιμοποιείται στην παρούσα διατριβή είναι η μέθοδος της υπερμοριακής προσέγγισης (Supermolecular Approach). Στην συνέχεια παρουσιάζονται αυτοί οι δύο τρόποι.

2.3.1.2 Η μέθοδος SAPT

Η μέθοδος SAPT χρησιμοποιείται για τον άμεσο υπολογισμό της διαμοριακής ενέργειας αλληλεπίδρασης και εφαρμόζεται στην κυματοσυνάρτηση του διμερούς. Η μέθοδος στηρίζεται στην συμμετρική Rayleigh-Schrodinger θεωρία διαταραχών⁸⁸ όπου η συνολική διαμοριακή ενέργεια υπολογίζεται από τους όρους όλων των τάξεων της διαταραχής. Η πρώτη τάξης ενέργεια αλληλεπίδρασης συμπεριλαμβάνει τις ενέργειες που οφείλονται στις ηλεκτροστατικές δυνάμεις Coulomb και ανταλλαγής-άπωσης. Η δεύτερης τάξης ενέργεια αλληλεπίδρασης συμπεριλαμβάνει τις ενέργειες που οφείλονται στις δυνάμεις επαγωγής και διασποράς^{89,90}. Αν και η μέθοδος θεωρείται από τις καλύτερες μεθόδους για τον άμεσο υπολογισμό της διαμοριακής ενέργειας παρουσιάζει πρόβλημα ακρίβειας στους ενεργειακούς όρους διαταραχής υψηλότερων τάξεων, καθώς και προβλήματα σύγκλισης⁹¹. Η ποσοτική και ποιοτική αποτελεσματικότητα της μεθόδου SAPT αποδείχτηκε ότι πρέπει να ελέγχεται από την



μέθοδο της υπερμοριακής προσέγγισης^{86,92}.

2.3.1.3 Η μέθοδος της υπερμοριακής προσέγγισης

Η μέθοδος της υπερμοριακής προσέγγισης (Supermolecular approach) είναι θεμελιωδώς η πιο απλή μέθοδος που μπορεί να εφαρμοστεί στον υπολογισμό της διαμοριακής ενέργειας αλληλεπίδρασης και στηρίζεται στον υπολογισμό των όρων της εξίσωσης (2.2) με συνήθεις ab-initio θεωρίες μοριακής κβαντικής μηχανικής (HF, MP2, CC, DFT, CI, MCSCF κλπ). Ο υπολογισμός των όρων που αφορούν τις ενεργειακές αλληλεπιδράσεις περισσότερων των δύο μορίων (n-body interactions) μπορεί να γίνει είτε με την χρήση της θεωρίας SAPT είτε με αναγωγή σε μικρότερους διμοριακούς υπολογισμούς^{93,94}.

Αναλυτικότερα, ο υπολογισμός της διαμοριακής ενέργειας αλληλεπίδρασης σε ένα διμερές μοριακό σύστημα υπολογίζεται σύμφωνα με την επόμενη εξίσωση:

$$\Delta E^{\text{int}} = E_{AB}(AB) - E_A(A) - E_B(B) \quad (2.4)$$

όπου $E_{AB}(AB)$ είναι ενέργεια του διμερούς μοριακού συστήματος και $E_A(A)$ και $E_B(B)$ οι ενέργειες των μονομερών. Στην εξίσωση (2.4) οι δείκτες συμβολίζουν το μοριακό σύστημα και τα σύμβολα μέσα στις παρενθέσεις το σύνολο βάσης που χρησιμοποιείται με την επιλεγμένη ab-initio μέθοδο. Η μοναδική προϋπόθεση που πρέπει να ισχύει για την εφαρμογή της μεθόδου είναι ότι: οι μέθοδοι κβαντικής μηχανικής που θα εφαρμοστούν για τον υπολογισμό των ενεργειών πρέπει να είναι συνεπής ως προς το μέγεθος (size-consistent). Αυτό σημαίνει ότι για μεγάλη διαμοριακή απόσταση R πρέπει να επαληθεύεται η εξίσωση:

$$E_{AB}(AB) = E_A(A) + E_B(B) \quad (2.5)$$

Με την χρήση ενός πλήρους συνόλου βάσης στους ab-initio κβαντικούς υπολογισμούς οι εξισώσεις (2.4) και (2.5) είναι ακριβής. Καθότι η χρήση πλήρους συνόλου βάσης στους μοριακούς υπολογισμούς προσομοίωσης δεν είναι εφικτή χρησιμοποιούμε σύνολα βάσης με πεπερασμένους όρους τα οποία περιγράφουν τις ατομικές συναρτήσεις των ατομικών τροχιακών. Η χρήση συναρτήσεων ατομικών τροχιακών με πεπερασμένους όρους, συνεπάγεται την εμφάνιση μίας ενεργειακής ποσότητας η οποία «καταργεί» την ισότητα της εξίσωσης (2.5) και αναλύεται στην επόμενη ενότητα.

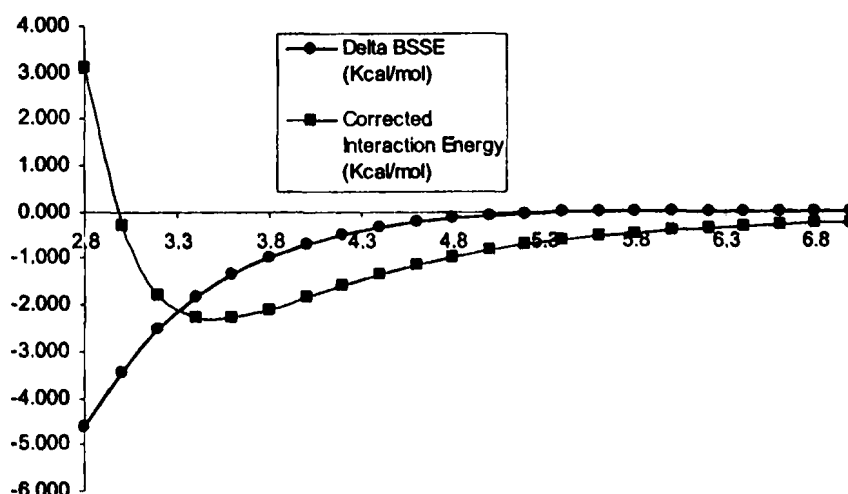


2.3.1.4. Το πρόβλημα του BSSE

Οι εξισώσεις (2.4) και (2.5) είναι ακριβείς είτε σε μεγάλη διαμοριακή απόσταση R είτε κάνοντας χρήση πλήρους συνόλου βάσης⁹⁵. Όταν η διαμοριακή απόσταση ελαττώνεται η εξίσωση (2.5) δεν ισχύει και η διαμοριακή ενέργεια αλληλεπίδρασης υπερεκτιμάται ποσοτικά^{93,96}. Το φαινόμενο αυτό οφείλεται στο γεγονός ότι τα επιμέρους μονομερή (του διμερούς) επηρεάζονται από το μεγαλύτερο σύνολο βάσης AB του διμερούς. Η υπερεκτίμηση της διαμοριακής ενέργειας αλληλεπίδρασης ονομάζεται «λάθος που προκύπτει από την υπέρθεση του συνόλου βάσεων» ή Basis Set Superposition Error (BSSE)⁹⁷. Αν και η ποσότητα αυτή αναδείχτηκε από τους Jansen και Ross⁹⁸ [115] το 1969 η ορολογία BSSE εδραιώθηκε το 1973.

Σε προσομοιώσεις διμερών μοριακών συστημάτων στοχεύουμε να υπολογίσουμε και να αφαιρέσουμε την ποσότητα αυτή από την διαμοριακή ενέργεια αλληλεπίδρασης. Αξίζει να σημειωθεί ότι το φαινόμενο του BSSE είναι και ενδομοριακό φαινόμενο, διότι σχετίζεται με το πλήθος των πεπερασμένων συναρτήσεων των ατομικών τροχιακών που εφαρμόζονται στους πρακτικούς *ab-initio* υπολογισμούς. Στην εικόνα 2.2 φαίνεται η ποσότητα BSSE συναρτήσει της διαμοριακής απόστασης R στο διμερές $C_6H_6-H_2O$. Ο υπολογισμός έγινε στο επίπεδο $MP2/6-31G^*$.

Εικόνα 2.2



Στην εικόνα αυτή φαίνεται η διακύμανση του BSSE και της διαμοριακής ενέργειας αλληλεπίδρασης συναρτήσει της διαμοριακής απόστασης σε υπολογισμό $MP2/6-31G^*$ στο διμερές $C_6H_6-H_2O$

Όπως φαίνεται στην εικόνα 2.2 η ποσότητα BSSE είναι αρκετά σημαντική σε μικρή διαμοριακή απόσταση ενώ αρχίζει να απαλείφεται μετά τα 5Å.

Η υπερμοριακή προσέγγιση είναι καλώς ορισμένη για τον υπολογισμό διαμοριακών ενεργειών αλληλεπίδρασης ειδικά σε διμερή συστήματα που συνδέονται με ασθενείς δυνάμεις van der Waals⁹⁶. Το κύριο μειονέκτημα της υπερμοριακής προσέγγισης είναι η εύρεση και η εφαρμογή της κατάλληλης μεθόδου ώστε να απαλοιφεί η ποσότητα BSSE. Υπάρχουν οι επόμενες υπολογιστικές διαδικασίες που εφαρμόζονται στην πράξη για να εξαλειφθεί η ποσότητα BSSE:

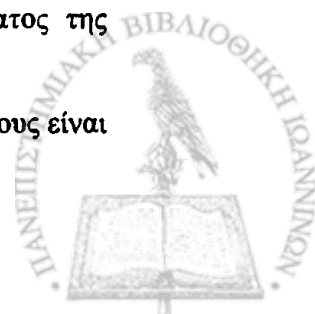
- Η μέθοδος counterpoise (CP)^{75,98}.
- Η μέθοδος της προσέγγισης της χημικής χαμιλτωνιανής (Chemical Hamiltonian Approach - CHA)⁹⁹.
- Η μέθοδος SCF-MI (Self Consistent Field for Molecular Interactions)^{100,101,102}.

Η μέθοδος counterpoise εφαρμόζεται εκ των υστέρων (a-posteriori) στους μοριακούς κβαντικούς υπολογισμούς ενώ η CHA και SCF-MI εφαρμόζονται εκ των προτέρων (a-priori) στην κβαντική κυματοσυνάρτηση του διμερούς. Η counterpoise μέθοδος αναφέρεται στην βιβλιογραφία ως μέθοδος των Boys Bernardi⁷⁵ και είναι η μέθοδος που εφαρμόζεται στην παρούσα διατριβή για τον υπολογισμό των διαμοριακών ενεργειών αλληλεπίδρασης ενώ αναλύεται στην επόμενη ενότητα.

Η μέθοδος CHA εφαρμόζεται στην μοριακή χαμιλτωνιανή με σκοπό να απαλλάξει τις συναρτήσεις των ατομικών τροχιακών από την ποσότητα BSSE. Αν και είναι μία πολύ καλή προσπάθεια για να εξαλείψει την φύση του προβλήματος έχει αρκετά μειονεκτήματα στην εφαρμογή της. Ο κύριος λόγος είναι ότι εφαρμόζεται σε σύνολα βάσεων τα οποία βελτιστοποιήθηκαν για να παράγουν μοριακές ιδιότητες (διπολική ροπή, πολωσιμότητα, ενέργεια σχηματισμού, κλπ) τις οποίες η CHA δεν μπορεί να αναπαράγει. Στην πράξη η μέθοδος CHA δίνει όμοια αποτελέσματα με την μέθοδο counterpoise, όμως η τελευταία είναι πιο γενική στην χρήσης της.

Η μέθοδος SCF-MI εφαρμόζεται στην θεωρία SCF ώστε κάθε μονομερές χρησιμοποιεί την δική του βάση συναρτήσεων ατομικών τροχιακών. Η μέθοδος αυτή συνήθως οδηγεί σε μία υπερδιόρθωση του BSSE λόγω έλλειψης όρων αλληλεπίδρασης στην συνολική κυματοσυνάρτηση. Αν και η μέθοδος SCF-MI είναι υλοποιημένη σε αρκετά πακέτα μοριακής κβαντικής μηχανικής όπως το GAMESS-US δεν χρησιμοποιείται ευρέως στην βιβλιογραφία λόγω του προβλήματος της υπερδιόρθωσης.

Ένα σημαντικό χαρακτηριστικό που αφορά τις τρεις προαναφερθέντες μεθόδους είναι



ότι μόνο οι μέθοδοι counterpoise και CHA οδηγούν στον μηδενισμό του BSSE όσο το σύνολο βάσης μεγαλώνει ενώ η μέθοδος SCF-MI δεν παρουσιάζει την ίδια ασυμπτωτική συμπεριφορά. Η ομοιότητα των μεθόδων counterpoise και CHA αποδεικνύεται⁹⁹ και από το γεγονός ότι όσο το σύνολο βάσης των ατομικών τροχιακών αυξάνεται, το BSSE ελαττώνεται γρηγορότερα απ' ότι στην πραγματικότητα από την διαμοριακή ενέργεια αλληλεπίδρασης. Οι δύο αυτές μέθοδοι παρουσιάζουν την ίδια ασυμπτωτική συμπεριφορά.

2.3.2. Η μέθοδος των Boys και Bernardi

Οι Boys and Bernardi πρότειναν έναν τρόπο για τον υπολογισμό της ποσότητας BSSE η οποία ονομάζεται counterpoise (CP) μέθοδος και εφαρμόζεται στην εξίσωση (2.4). Στην CP μέθοδο γίνονται επιπλέον υπολογισμοί εξισώνοντας τις συναρτήσεις των συνόλων βάσεων A και B των μονομερών με το σύνολο βάσης του διμερούς. Αυτό πραγματοποιείται με το να υπολογίσουμε την ολική ηλεκτρονική ενέργεια των μονομερών με το σύνολο βάσης AB του διμερούς. Η μέθοδος CP είναι ευρέως διαδεδομένη και χρησιμοποιείται σε όλα σχεδόν τα επίπεδα *ab-initio* κβαντικών υπολογισμών. Πολλές μελέτες έδειξαν ότι:

- Η ποσότητα BSSE απαλείφεται όπως και με την «εκ των προτέρων» μέθοδο CHA σε διαφορετικά επίπεδα θεωρίας.
- Τα αποτελέσματα είναι συγκρίσιμα με την απαλλαγμένη από το BSSE μέθοδο SAPT.
- Ασυμπτωτικά τείνει στο μηδέν εφόσον μεγαλώνουν τα σύνολα βάσεων που περιγράφουν τα ατομικά τροχιακά⁹⁵.
- Χρησιμοποιώντας μικρά σύνολα βάσεων δίνει πολύ καλό ποσοτικό προσδιορισμό του BSSE. Έτσι μπορεί να εφαρμοστεί ευκολότερα σε μεγάλα μοριακά συστήματα βιολογικής σημασίας όπου συνήθως η χρήση μεγάλων συνόλων βάσεων είναι απαγορευτική¹⁰³.

Η μέθοδος CP εφαρμόζεται εκ των υστέρων (*a-posteriori*), δηλαδή εφόσον γίνουν οι υπολογισμοί της εξίσωσης (2.4). Στην πράξη στους κβαντικούς μοριακούς υπολογισμούς χρησιμοποιούνται τα άτομα «φαντάσματα» (*ghost atoms*) για να υπολογίσουμε την ποσότητα BSSE. Αυτό γίνεται ορίζοντας τα άτομα του μονομερούς να έχουν μηδενικό φορτίο αλλά να συμπεριλαμβάνονται οι συναρτήσεις των ατομικών τροχιακών τους. Ως συνέπεια αυτού, εκτός από τους τρεις υπολογισμούς



της εξίσωσης (2.4) χρειάζονται δύο επιπλέον υπολογισμοί. Ο ένας υπολογισμός γίνεται για το μονομερές A στην συνολική βάση AB του διμερούς και ο άλλος για το μονομερές B στην συνολική βάση AB.

Η επόμενη εξίσωση δίνει την ζητούμενη ποσότητα BSSE:

$$\delta^{BSSE} = E_A(AB) - E_A(A) + E_B(AB) - E_B(B) \quad (2.6)$$

Αναλυτικότερα ο υπολογισμός $E_A(AB)$ σημαίνει ότι ορίζουμε τα άτομα του μονομερούς B ως "ghost". Αυτό πραγματοποιείται θέτοντας το φορτίο των ατόμων του μονομερούς B ίσο με μηδέν και χρησιμοποιώντας για τον υπολογισμό της ολικής ηλεκτρονικής ενέργειας το σύνολο βάσης AB του διμερούς. Η ενέργεια που υπολογίζεται με αυτόν τον τρόπο για το μονομερές A είναι χαμηλότερη από την ενέργεια $E_A(A)$ λόγω του ότι το σύνολο βάσης AB είναι περισσότερο πλήρες για το μονομερές. Αυτός είναι και ο κύριος λόγος που η χρήση πλήρους συνόλου βάσεως για υπολογισμούς στην εξίσωση (2.4) είναι απαλλαγμένη από την ποσότητα BSSE. Ακριβώς η ίδια διαδικασία χρησιμοποιείται και για τον υπολογισμό της ποσότητας BSSE για το μονομερές B. Η συνολική ποσότητα που υπολογίζεται με την εξίσωση (2.6) ελαττώνει την συνολική διαμοριακή ενέργεια αλληλεπίδρασης.

Κάνοντας χρήση των εξισώσεων (2.4) και (2.6) προκύπτει η επόμενη εξίσωση:

$$E^{CP} = E_{AB}(AB) - \delta^{BSSE} \quad (2.7)$$

Η εξίσωση (2.7) δίνει την ολική ηλεκτρονική ενέργεια του διμερούς μοριακού συστήματος απαλλαγμένη από το BSSE. Η διορθωμένη διαμοριακή ενέργεια του μοριακού συστήματος υπολογίζεται από την εξίσωση:

$$\Delta E^{CP} = E_{AB}(AB) - E_A(AB) - E_B(AB) \quad (2.8)$$

Για τον υπολογισμό διαμοριακών ενεργειών αλληλεπίδρασης στην υπερεπιφάνεια δυναμικής ενέργειας ή σε μία περιοχή αυτής απαιτείται να γίνει γεωμετρική αριστοποίηση του μοριακού διμερούς ελαχιστοποιώντας την ολική ηλεκτρονική ενέργεια με παραγώγους ως προς την γεωμετρία. Η διαδικασία αυτή είναι κοινή σε όλους του κβαντικούς υπολογισμούς μοριακών συστημάτων. Στην βελτιστοποιημένη γεωμετρία γίνονται οι επιπλέον τέσσερις υπολογισμοί της εξίσωσης (2.6). Στο πλήθος των πακέτων κβαντικής μοριακής προσομοίωσης οι επιπλέον υπολογισμοί γίνονται από τον χρήστη φτιάχνοντας τα κατάλληλα αρχεία εισόδου για τα πακέτα προσομοίωσης.

Στο σύστημα MERGAM αυτοματοποιήσαμε την διαδικασία αυτή και με ένα αρχείο εισόδου τύπου GAMESS-US γίνονται όλοι οι επιμέρους υπολογισμοί και εξάγονται



όλες οι ενεργειακές ποσότητες της εξίσωσης (2.6). Στον πίνακα 2.1 φαίνεται το αποτέλεσμα που δίνει το σύστημα MERCAM σε έναν υπολογισμό στο διμερές $C_6H_6-H_2O$.

Πίνακας 2.1

Interaction Energy:	-1.436 Kcal/mol
CP Corrected Int. Energy:	-1.209 Kcal/mol
DELTA BSSE:	-0.227 Kcal/mol
DELTA BSSE (au):	-0.00036251
Total Energy (au):	-306.73761846
Total Corrected Energy (au):	-306.73725596

Αποτελέσματα υπολογισμού σε διμερές με όλες τις τιμές των εξισώσεων (2.1)-(2.5)

2.3.2.1 Αναλυτικές παράγωγοι των CP-διορθωμένων ενεργειών

Μία άλλη τεχνική που υλοποιήθηκε στο σύστημα MERCAM αφορά την γεωμετρική αριστοποίηση ενός διμερούς μοριακού συστήματος ελαχιστοποιώντας την BSSE-διορθωμένη ολική ηλεκτρονική ενέργεια E^{CP} με χρήση αναλυτικών παραγώγων¹⁰⁴. Η διαδικασία αυτή είναι πολύ αποτελεσματική στην μελέτη, με μεγάλη ακρίβεια, μιας περιοχής της υπερεπιφάνειας διαμοριακής δυναμικής ενέργειας. Εξ' ορισμού το πακέτο MERLIN/MCL παρέχει μεθόδους βελτιστοποίησης με αναλυτικές και αριθμητικές παραγώγους. Η χρήση αριθμητικών παραγώγων εφαρμόζεται σε περιπτώσεις που η αντικειμενική συνάρτηση δεν έχει αναλυτικές παραγώγους. Στην περίπτωση της μεθόδου CP ενσωματώσαμε στο σύστημα MERCAM αναλυτικές παραγώγους της ενέργειας που περιγράφεται από την εξίσωση (2.7) ως προς την γεωμετρία. Η επόμενη εξίσωση δίνει τον υπολογισμό των αναλυτικών παραγώγων της εξίσωσης (2.7):

$$\frac{\partial E^{CP}}{\partial x_j} = \frac{\partial E_{AB}(AB)}{\partial x_j} - \frac{\partial E_A(AB)}{\partial x_j} + \frac{\partial E_A(A)}{\partial x_j} - \frac{\partial E_B(AB)}{\partial x_j} + \frac{\partial E_B(B)}{\partial x_j} \quad (2.9)$$

Όπου x_j είναι η j -γεωμετρική παράμετρος του μοριακού συστήματος σε καρτεσιανές συντεταγμένες. Σε κάθε γεωμετρική διεύθυνση, κατά την διάρκεια την ενεργειακής μοριακής ελαχιστοποίησης, το σύστημα MERCAM εκτελεί όλους τους υπολογισμούς της εξίσωσης (2.9) αυτόματα όπως θα δούμε και στις επόμενες ενότητες των εφαρμογών του συστήματος.



2.4. Η υλοποίηση της μεθόδου των Boys και Bernardi στο MERGAM.

Το πακέτο MERLIN/MCL παρέχει μία εσωτερική βιβλιοθήκη ρουτινών με τις οποίες ο χρήστης μπορεί να επεκτείνει τις λειτουργίες του MERLIN και να προγραμματίσει δικούς του αλγόριθμους και λειτουργικές υπομονάδες⁷⁹. Οι επεκτάσεις αυτές, που μπορεί να προγραμματίσει ο χρήστης για να επεκτείνει το MERLIN, αναφέρονται με την κωδικοποιημένη ονομασία "plugins" και δεν είναι τίποτα άλλο εκτός από υπορουτίνες σε Fortran 77. Στο σύστημα MERGAM αναπτύχθηκαν οι τρεις επεκτάσεις GMSGRAD, CPC και CPCOPT όπως φαίνονται και στην εικόνα 2.1. Ο πηγαίος κώδικας των επεκτάσεων βρίσκεται στο αρχείο "plug.d". Η ονομασία των αντίστοιχων Fortran 77 ρουτινών είναι PLUG1, PLUG2 και PLUG3. Οι τρεις αυτές ρουτίνες καλούν C-wrapper συναρτήσεις καθώς και ρουτίνες της εσωτερικής βιβλιοθήκης του MERLIN. Η περιγραφή των νέων εντολών που αναπτύχθηκαν ως επεκτάσεις ακολουθεί αναλυτικά:

- **GMSGRAD:** Η εντολή αυτή αντιστοιχεί στην υπορουτίνα PLUG1 και η λειτουργία της είναι να ενεργοποιεί την καθολική μεταβλητή GRADOPT του συστήματος MERGAM. Όταν είναι ενεργοποιημένη η μεταβλητή αυτή το MERGAM ζητάει από το GAMESS-US να εκτελέσει μοριακό κβαντικό υπολογισμό με πρώτες αναλυτικές παραγώγους ακόμα και αν ζητείται από το MERLIN μόνο η τιμή της αντικειμενικής συνάρτησης σε οποιαδήποτε γεωμετρική διεύθυνση. Η λειτουργία αυτή εμπεριέχει κλήση στην C-wrapper συνάρτηση GMS_OPTTYP.
- **CPC:** Η εντολή αυτή αντιστοιχεί στην υπορουτίνα PLUG2 και αφορά τον υπολογισμό διαμοριακών ενεργειών αλληλεπίδρασης διορθωμένων ως προς την ποσότητα BSSE σε διμερή μοριακά συστήματα. Η εντολή αυτή προγραμματίστηκε ώστε να υπολογίζονται αυτόματα οι ποσότητες των εξισώσεων (2.1)-(2.4). Εκτός από την τιμή της αντικειμενικής συνάρτησης που υπολογίζεται με κλήση στην FUNMIN, οι τέσσερις επιπλέον υπολογισμοί γίνονται με αντίστοιχες εκτελέσεις του πακέτου GAMESS-US. Η εντολή CPC απαιτεί 5 παραμέτρους όπως φαίνονται στον πίνακα 2.2. Ο πίνακας 2.2 παρουσιάζει ακριβώς την μορφή του "panel" που θα συναντήσει κάποιος εκτελώντας την εντολή στο περιβάλλον του MERLIN. Η C-wrapper συνάρτηση που καλείται είναι η GMS_CP.
- **CPCOPT:** Η εντολή αυτή αντιστοιχεί στην υπορουτίνα PLUG3 και χρησιμοποιείται για να ενεργοποιήσει γεωμετρική βελτιστοποίηση με διορθωμένη ενέργεια ως προς την ποσότητα BSSE. Απαιτεί τις ίδιες παραμέτρους που φαίνονται στον πίνακα 2.2. Σύμφωνα με την εικόνα 2.1 της αρχιτεκτονικής ενεργοποιεί την καθολική



μεταβλητή GCPCOPT καλώντας την C-wrapper συνάρτηση GMS_CPCOPT και αναθέτει στις καθολικές μεταβλητές ST1, HW1, ST2 και HW2 τις παραμέτρους εισόδου του πίνακα 2.2. Οι μεταβλητές αυτές χρησιμοποιούνται για τον προσδιορισμό των ατόμων από τα οποία αποτελείται κάθε μονομερές.

Πίνακας 2.2 (Merlin Command Panel)

Ind	Keyword	Description	Value	Allowed values
1)	MAF	Monomer A First Atom	1	Any int \geq 1
2)	MAN	Monomer A Number of Atoms	12	Any int \geq 1
3)	MBF	Monomer B First Atom	13	Any int \geq 1
4)	MBN	Monomer B Number of Atoms	3	Any int \geq 1
5)	CENERGY	non corrected/corrected...	1	{0,1}
6)	CANCEL	Cancel / Proceed	1	{0,1}

MAF: Το πρώτο άτομο του πρώτου μονομερούς; **MAN:** Ο συνολικός αριθμός των ατόμων που αποτελούν το πρώτο μονομερές; **MBF:** Το πρώτο άτομο του δεύτερου μονομερούς; **MBN:** Ο συνολικός αριθμός των ατόμων που αποτελούν το δεύτερο μονομερές; **CENERGY:** Μεταβλητή που καθορίζει τον τύπο της ενέργειας που θα μας επιστρέψει ο υπολογισμός στο GAMESS-US; **CANCEL:** Ακυρώνει ή εκτελεί την διαδικασία.



2. 5. Εφαρμογές με το σύστημα MERGAM

Η λειτουργικότητα και αποτελεσματικότητα του συστήματος MERGAM ελέγχθηκε με την μελέτη των διμερών μοριακών συστημάτων HF-H₂O and C₆H₆-H₂O. Τα διμερή αυτά επιλεχθήκαν ως αντιπροσωπευτικά παραδείγματα ενός ισχυρού και ενός ασθενούς δεσμού υδρογόνου ενώ παρουσιάζουν ιδιότητες με ιδιαίτερο ενδιαφέρον^{104,105,106}. Εντελώς διαφορετικές υπολογιστικές προσεγγίσεις εφαρμόστηκαν σε κάθε διμερές, ώστε να γίνει εξαντλητική μελέτη της λειτουργικότητας του συστήματος MERGAM και να εξεταστεί η αξιοπιστία του. Η μελέτη στο διμερές C₆H₆-H₂O γίνεται για πρώτη φορά στο επίπεδο που παρουσιάζεται σε αυτή την διατριβή και αφορά την αυτοματοποιημένη διερεύνηση μιας περιοχής της διαμοριακής επιφάνειας δυναμικής ενέργειας.

2.5.1. CP-διορθωμένη γεωμετρική βελτιστοποίηση στο διμερές HF-H₂O

Το αίτημα για υπολογισμούς υπερεπιφανειών CP-διορθωμένης διαμοριακής ενέργειας έχει εντοπιστεί στην βιβλιογραφία^{93,107,108} και έχει ζητηθεί αυτοματοποιημένη μέθοδος γεωμετρικής αριστοποίησης με αναλυτικές παραγώγους.

Στο σύστημα MERGAM ενσωματώθηκε η δυνατότητα αυτή επεκτείνοντας τις δυνατότητες του πακέτου GAMESS-US το οποίο δεν παρέχει τέτοια αυτοματοποιημένη μέθοδο. Στην πράξη προγραμματίστηκε η εξίσωση (2.9). Για να ελεγχθεί η αξιοπιστία της αυτοματοποιημένης αυτής μεθόδου κάναμε γεωμετρική βελτιστοποίηση στο διμερές HF-H₂O. Ο κύριος στόχος της εφαρμογής αυτής είναι ο έλεγχος της εγκυρότητας των αποτελεσμάτων σε σύγκριση με προηγούμενη βιβλιογραφική μελέτη¹⁰⁴. Η είσοδος στο MERLIN για τον υπολογισμό αυτό έχει ως εξής:

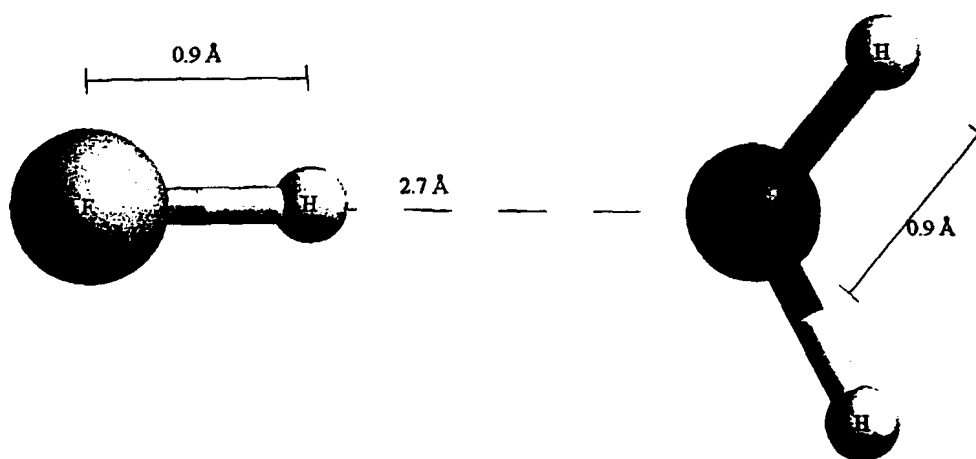
```
anal
gmsgrad
eval
tolmin acc 1.e-4
cpc MAF 1 MAN 2 MBF 3 MBN 3
cpcopt MAF 1 MAN 2 MBF 3 MBN 3
noeval
init NPL 1
eval
tolmin acc 1.e-4
cpc
```

Στην παραπάνω αλληλουχία εντολών φαίνεται πόσο σημαντικό είναι ένα εξελιγμένο



περιβάλλον στρατηγικής βελτιστοποίησης σε μοριακούς κβαντικούς υπολογισμούς όπως το MERGAM. Η στρατηγική αυτή για το διμερές HF-H₂O εκτελεί δύο διαδοχικές γεωμετρικές βελτιστοποιήσεις. Η πρώτη είναι η κλασική περίπτωση ενώ η δεύτερη είναι CP-διορθωμένη. Τα αποτελέσματα της στρατηγικής εξάγονται σε ένα αρχείο και είναι σε πλήρη συμφωνία με προηγούμενη μελέτη¹⁰⁴ (-176.049751 a.u. για την μη-διορθωμένη και -176.048211 a.u. για την CP-διορθωμένη γεωμετρική βελτιστοποίηση).

Εικόνα 2.3



Η CP-διορθωμένη βελτιστοποιημένη γεωμετρία του HF-H₂O

Στην εικόνα 2.3 φαίνεται η βελτιστοποιημένη γεωμετρία του διμερούς η οποία προέκυψε με αναλυτικές CP-διορθωμένες πρώτες παραγώγους.

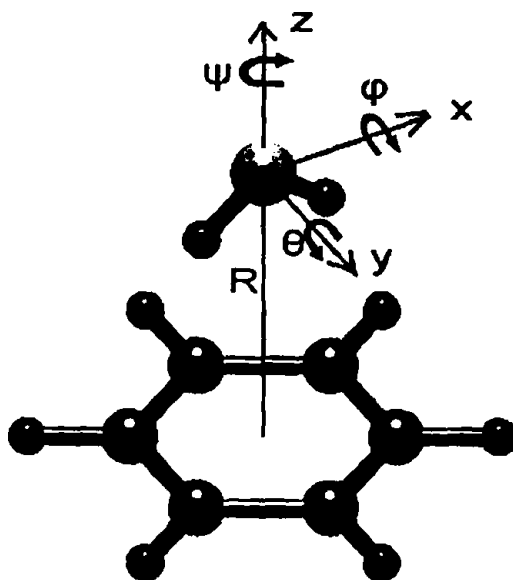
Οι υπολογισμοί έγιναν στο σύνολο βάσης RHF/6-31G** και εφαρμόστηκε ο αλγόριθμος βελτιστοποίησης Tolmin. Αν και το επίπεδο υπολογισμών δεν περιλαμβάνει μεθόδους που να λαμβάνουν υπ' όψη την δυναμική συσχέτιση της ηλεκτρονικής ενέργειας είναι κοντά σε πειραματική μελέτη¹⁰⁹, όπου βρέθηκε η απόσταση F-O ίση με 2.62Å.

2.5.2. Εξερεύνηση μιας περιοχής της CP-διορθωμένης διαμοριακής υπερεπιφάνειας δυναμικής ενέργειας του διμερούς $C_6H_6-H_2O$

Η μελέτη μιας περιοχής της CP-διορθωμένης υπερεπιφάνειας διαμοριακής ενέργειας του διμερούς $C_6H_6-H_2O$ υλοποιήθηκε με σκοπό να διασαφηνίσει την αποτελεσματικότητα μιας στρατηγικής διαδικασίας που μπορεί να επιτευχθεί με την χρήση του MERGAM. Για τους υπολογισμούς χρησιμοποιήθηκε η μέθοδος MP2/6-31G** η οποία λαμβάνει υπ' όψη την δυναμική ηλεκτρονική συσχέτιση. Η επέκταση (plug-in) CPC που αναπτύχθηκε στο MERGAM χρησιμοποιήθηκε για τον υπολογισμό των CP-διορθωμένων ενεργειών του διμερούς μετά από κάθε γεωμετρική βελτιστοποίηση.

Το αρχικό στάδιο της μελέτης αυτής αφορούσε την γεωμετρική αριστοποίηση του κάθε μονομερούς ξεχωριστά στο επίπεδο θεωρίας MP2/6-31G**. Στη συνέχεια συγκροτήθηκε το διμερές με τις βελτιστοποιημένες γεωμετρίες ώστε να υποβληθεί στην στρατηγική μελέτη της δυναμικής υπερεπιφάνειας που σχεδιάστηκε. Η αρχική γεωμετρία φαίνεται στην εικόνα 2.4 όπου τα άτομα υδρογόνου του H_2O κατευθύνονται προς το επίπεδο του C_6H_6 .

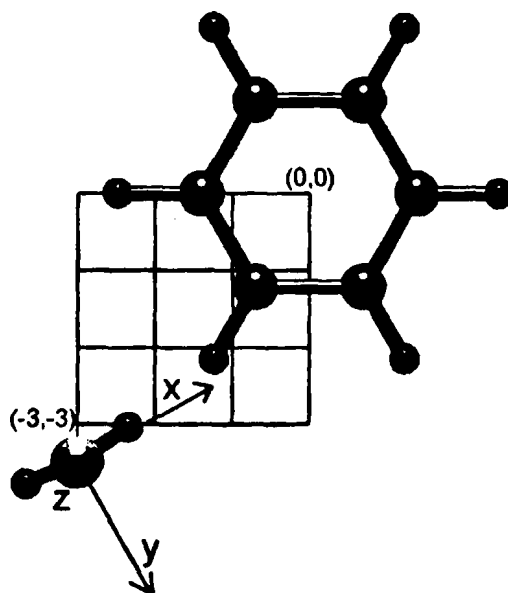
Εικόνα 2.4



Το σύστημα συντεταγμένων στην αριστοποιημένη MP2/6-31G** γεωμετρία του διμερούς. Οι σχετικές γωνίες Euler ορίζονται να είναι: $\delta\phi=0^\circ$, $\delta\theta=0^\circ$ and $\delta\psi=0^\circ$ ως προς τους άξονες x-, y- και z αντίστοιχα. Το άτομο του οξυγόνου βρίσκεται σε διαμοριακή απόσταση $R=3.5\text{\AA}$ πάνω από το κέντρο μάζας του βενζολίου.

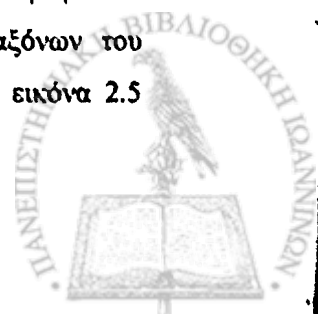
Στην στρατηγική που εφαρμόστηκε το C_6H_6 θεωρείται άκαμπτο σύστημα, ενώ μόνο τα υδρογόνα του H_2O επιτρέπεται να μεταβάλλονται σε όλους τους βαθμούς ελευθερίας τους. Κάθε μονομερές κείται στο δικό του καρτεσιανό σύστημα συντεταγμένων. Το άτομο οξυγόνου ορίζει την αρχή των αξόνων για το καρτεσιανό σύστημα συντεταγμένων του H_2O , ενώ ως αρχή των αξόνων για το C_6H_6 λαμβάνεται το κέντρο μάζας του. Ο σχετικός προσανατολισμός του H_2O ως προς το C_6H_6 , όπως φαίνεται στην εικόνα 2.4 ορίζεται από α) την απόσταση R μεταξύ της αρχής των αξόνων των δύο μονομερών και β) τις τρεις γωνίες Euler φ , θ και ψ του συστήματος συντεταγμένων του H_2O . Οι γωνίες Euler φ , θ και ψ στον 3-διάστατο καρτεσιανό χώρο αντιστοιχούν στην περιστροφή ενός αντικειμένου γύρω από τους τρεις άξονες x , y , z και χρησιμοποιούνται για να περιγράψουν το νέο προσανατολισμό ενός αντικειμένου που υπόκειται περιστροφή γύρω από την αρχή των αξόνων.

Εικόνα 2.5



Το 4x4 πλέγμα που χρησιμοποιήθηκε για την στρατηγική διαδικασία μελέτης περιοχής της επιφάνειας διαμοριακής ενέργειας στο διμερές $C_6H_6-H_2O$.

Εφόσον ορίστηκε το πρόβλημα που θα μελετηθεί, η στρατηγική διαδικασία (παράρτημα Ε) αναπτύχθηκε στην γλώσσα MCL του MERLIN και εφαρμόστηκε. Για την διαδικασία αυτή θεωρούμε ότι στο επίπεδο του C_6H_6 υπάρχει ένα τετραγωνικό πλέγμα του οποίου το σημείο $(0,0)$ συμπίπτει με την αρχή των αξόνων του συστήματος συντεταγμένων του C_6H_6 . Το πλέγμα όπως φαίνεται στην εικόνα 2.5



αποτελείται από δεκαέξι ισαπέχοντες κόμβους με απόσταση 1Å μεταξύ τους. Οι διακριτές συνταγμένες των κόμβων του πλέγματος ορίζονται από το σύνολο $[-3,0] \times [-3,0] \text{Å}$. Σε κάθε βήμα της στρατηγικής διαδικασίας κινείται το άτομο του οξυγόνου του H_2O στους κόμβους του πλέγματος αλλά σε ύψος 3.5Å (η z-συντεταγμένη του O ισούται με 3.5). Με αυτό τον τρόπο καλύπτεται η περιοχή της δυναμική διαμοριακής υπερεπιφάνειας που μελετάται λαμβάνοντας υπ' όψη την συμμετρία C_{2v} του βενζολίου (εικόνα 2.6). Σε κάθε σημείο γίνεται γεωμετρική αριστοποίηση του συστήματος με τους επιτρεπτούς βαθμούς ελευθερίας που προαναφέρθηκαν και μετά υπολογίζεται αυτόματα η ποσότητα BSSE. Με την ολοκλήρωση της αυτοματοποιημένης στρατηγικής διαδικασίας λαμβάνονται όλα τα γεωμετρικά και CP-διορθωμένα ενεργειακά στοιχεία της συγκεκριμένης περιοχής της διαμοριακής δυναμικής υπερεπιφάνειας.

Πίνακας 2.3

Σημείο	Συντεταγμένες Οξυγόνου ^a			Γωνίες Euler ^b			Διαμοριακή Ενέργεια ^c
	X	Y	Z	$\delta\phi$	$\delta\theta$	$\delta\psi$	ΔE^{CP} MP2/6-31G**
1	-3	-3	3.5	0.138	-0.011	0.000	-0.27
2	-3	-2	3.5	0.074	0.020	0.000	-0.67
3	-3	-1	3.5	-0.302	-0.094	0.000	-0.93
4	-3	0	3.5	-25.380	-4.333	0.976	-1.07
5	-2	-3	3.5	16.330	10.536	1.516	-0.69
6	-2	-2	3.5	8.231	10.542	0.761	-1.16
7	-2	-1	3.5	0.020	-0.048	0.000	-1.33
8	-2	0	3.5	-12.080	-2.830	0.299	-1.41
9	-1	-3	3.5	22.691	17.337	3.504	-0.98
10	-1	-2	3.5	17.061	16.067	2.426	-1.38
11	-1	-1	3.5	0.176	0.092	0.000	-1.54
12	-1	0	3.5	-0.128	-0.135	0.000	-1.81
13	0	-3	3.5	52.413	49.950	25.825	-1.40
14	0	-2	3.5	22.873	14.126	2.872	-1.51
15	0	-1	3.5	12.819	6.706	0.754	-1.97
16	0	0	3.5	0.000	0.000	0.000	-2.33

^aΟι συντεταγμένες του οξυγόνου σε Å. ^bΟι γωνίες Euler υπολογίζονται σε μοίρες (deg). ^cΗ διαμοριακή ενέργειαν αλληλεπίδρασης υπολογίζεται σε Kcal/mol.

Τα γεωμετρικά χαρακτηριστικά και ΔE^{CP} (MP2/6-31G**) ενέργεια όπως υπολογίστηκαν στα δεκαέξι σημεία του πλέγματος από την αυτοματοποιημένη στρατηγική βελτιστοποίησης.

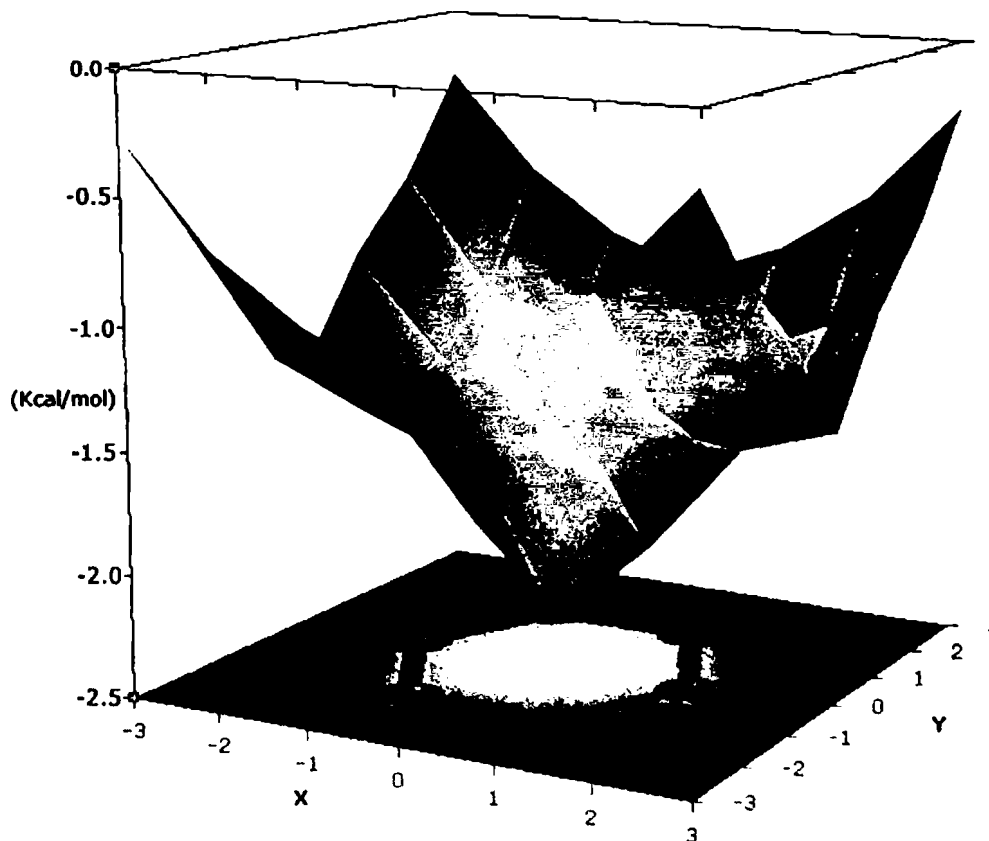
Χρησιμοποιώντας τα ενεργειακά αποτελέσματα της στρατηγικής διαδικασίας που εφαρμόσαμε με το MERGAM κατασκευάστηκε η πλήρης διαμοριακή επιφάνεια



δυναμικής ενέργειας που απεικονίζεται στην εικόνα 2.6. Το αποτέλεσμα των υπολογισμών φαίνεται στον πίνακα 2.3. Η διαμοριακή επιφάνεια δυναμικής ενέργειας βρίσκεται στην περιοχή $\Delta\Delta E^{CP} = \Delta E^{CP}_{\min} - \Delta E^{CP}_{\max} = 2.06$ Kcal/mol. Οι βελτιστοποιημένες διευθετήσεις των ατόμων υδρογόνου του νερού μετρήθηκαν με τις διαφορές των γωνιών Euler, του συστήματος συντεταγμένων του νερού, από την αρχική τους θέση που φαίνεται στην εικόνα 2.4.

Ο πίνακας 2.3 περιέχει αναλυτικά τα γεωμετρικά και ενεργειακά χαρακτηριστικά της CP-διορθωμένης επιφάνειας δυναμικής ενέργειας. Παρατηρώντας τις διακυμάνσεις των γωνιών Euler φαίνεται ότι μεγάλες διαφορές των γωνιών αντιστοιχούν σε πολύ μικρές ενεργειακές διαφορές (σημείο 13 στον πίνακα 2.3).

Εικόνα 2.6



Η επιφάνεια δυναμικής διαμοριακής ενέργειας όπως υπολογίστηκε σε απόσταση 3.5Å μεταξύ των μονομερών.

Οι μεγαλύτερες αποκλίσεις παρατηρούνται στις θέσεις που το νερό βρίσκεται πάνω από δεσμούς C-C και C-H γεγονός που οφείλεται στους ασθενείς δεσμούς υδρογόνου

OH-φ, όπου φ είναι ο αρωματικός δακτύλιος^{110, 111, 112}.

Σύμφωνα με προηγούμενες μελέτες η διαμοριακή ενέργεια αλληλεπίδρασης ΔE^{CP} του διμερούς $C_6H_6-H_2O$ έχει υπολογιστεί από -1.4 έως -2.8 Kcal/mol. Από την στρατηγική διαδικασία που εφαρμόστηκε αναδεικνύεται ως ελάχιστο το σημείο 16 του πίνακα 2.3. Η CP-διορθωμένη ενέργεια στο σημείο αυτό υπολογίστηκε να είναι $\Delta E^{CP} = 2.33$ Kcal/mol. Η τιμή αυτή είναι σε πολύ καλή συμφωνία με προηγούμενες θεωρητικές και πειραματικές μελέτες^{70,105,106,111} της διαμοριακής ενέργειας αλληλεπίδρασης.

2.5.2.1. Υπολογιστικό κόστος της διερεύνησης

Ο κώδικας που αφορά την διερεύνηση μιας περιοχής της επιφάνειας δυναμικής ενέργειας του διμερούς $C_6H_6-H_2O$ και δίνεται στο παράρτημα Ε, εκτελέστηκε 4 φορές μεταξύ των οποίων και παράλληλα. Τα αποτελέσματα δίνονται στο πίνακα 2.4

Πίνακας 2.4

A/A	CPUs	GAMESS-US CALLED BY MCL	Time (sec)
1	1xIntel Pentium III 800MHz	794	712450
2	1xIntel Core 2 Duo 2.1GHz	794	154027
3	2xIntel Core 2 Duo 2.1GHz (parallel)	794	67308
4	4xIntel Core 2 Duo 2.1GHz (parallel)	794	38021

2.6. Τεχνικές οδηγίες εκτέλεσης υπολογισμών χρησιμοποιώντας το MERCAM

Στο παράρτημα Γ δίνονται οι οδηγίες εγκατάστασης του συστήματος MERCAM. Στο παράρτημα Δ φαίνεται το αρχείο εισόδου (GAMESS-US input file), η αλληλουχία εντολών και η έξοδος από το σύστημα ως παράδειγμα ελέγχου. Στο σύστημα υπάρχει on-line βοήθεια για όλες τις εντολές και όλες τις επεκτάσεις (plug-in). Το αρχείο εισόδου πρέπει να δίνεται σε καρτεσιανές συντεταγμένες και ο τύπος υπολογισμού πρέπει να είναι ενεργειακός (RUNTYP=ENERGY). Εφόσον χρειαστεί το σύστημα εκτελεί αυτόματα τους υπολογισμούς πρώτων και δεύτερων παραγώγων. Όλος ο κώδικας της βιβλιοθήκης διασύνδεσης βρίσκεται στο παράρτημα ΣΤ.



2.7. Συμπεράσματα.

Το κεφάλαιο αυτό αναφέρεται στην δημιουργία ενός προγραμματιζόμενου συστήματος βελτιστοποίησης διασυνδέοντας τα πακέτα GAMESS-US και MERLIN/MCL. Για τον σκοπό αυτό γίνεται αναφορά στα πακέτα GAMESS-US και MERLIN/MCL. Επίσης γίνεται μια αναφορά στην θεωρία των διαμοριακών ενεργειών αλληλεπίδρασης τις οποίες το διασυνδεδεμένο σύστημα μπορεί να υπολογίσει με *ab-initio* μεθόδους μοριακής μηχανικής εφαρμόζοντας την μέθοδο *counterpoise*. Το σύστημα MERGAM προσθέτει επιπλέον ευελιξία στο πακέτο GAMESS-US για υπολογισμούς κβαντικής χημείας. Με τη χρήση της γλώσσας υψηλού επιπέδου MCL μπορούν να αναπτυχθούν ευφυής στρατηγικές για την μελέτη πολύπλοκων μοριακών δομών. Επίσης το σύστημα παρέχει μια ποικιλία από αλγόριθμους βελτιστοποίησης οι οποίοι μπορούν να εφαρμοστούν στην μελέτη επιφανειών δυναμικής ενέργειας μοριακών συστημάτων. Το περιβάλλον που αναπτύχθηκε προσφέρει αυτόματη CP-διορθωμένη γεωμετρική βελτιστοποίηση των διαμοριακών επιφανειών της ενέργειας αλληλεπίδρασης διμερών μοριακών συστημάτων με χρήση πρώτων αναλυτικών παραγώγων.

Το ενοποιημένο σύστημα ελέγχθηκε και εκτελείται σε όλες τις υπολογιστικές αρχιτεκτονικές με λειτουργικό σύστημα UNIX και Linux.

Η εφαρμογή του συστήματος στον υπολογισμό της διαμοριακής ενέργειας του διμερούς HF-H₂O παρουσιάζει την διαδικασία μιας CP-διορθωμένης γεωμετρικής βελτιστοποίησης η οποία είναι πολύ χρήσιμη για παρόμοιες μελέτες σε διμερή.

Η εφαρμογή του MERGAM στο διμερές C₆H₆-H₂O στοχεύει στην ανακατασκευή και εξερεύνηση μιας μεγάλης περιοχής της διαμοριακής ενέργειας αλληλεπίδρασης. Για την μελέτη αυτή εφαρμόστηκε για πρώτη φορά μία μεθοδολογία με την μορφή στρατηγικής διαδικασίας κάνοντας χρήση της γλώσσας MCL που παρέχει το MERLIN. Η μεθοδολογία αυτή μπορεί να επεκταθεί σε μια μεγάλη οικογένεια από διμερή για την μελέτη των διαμοριακών ενεργειών αλληλεπίδρασης που οφείλονται σε ασθενείς δεσμούς υδρογόνου και δυνάμεις van der Waals.

Το σύστημα MERGAM μπορεί επίσης να χρησιμοποιηθεί για περαιτέρω μελέτες που αφορούν πολύπλοκα διαμοριακά προβλήματα τα οποία σχετίζονται με τα γεωμετρικά και ενεργειακά χαρακτηριστικά ελκτικών αλληλεπιδράσεων μεταξύ διμερών που περιέχουν αρωματικούς δακτυλίους της μορφής XH..π, όπου X=O,N,S,... και π ο αρωματικός δακτύλιος.

Η τεχνολογία που χρησιμοποιήθηκε για την ανάπτυξη του συστήματος MERGAM



είναι δυνατόν να χρησιμοποιηθεί για την διασύνδεση και άλλων πακέτων ερευνητικού ενδιαφέροντος με το GAMESS-US ή με το MERLIN. Η τεχνολογική αυτή δυνατότητα διασύνδεσης που παρέχει το MERGAM εφαρμόστηκε για την μελέτη που γίνεται το κεφάλαιο 4.



ΚΕΦΑΛΑΙΟ 3

**Υπολογιστική προσομοίωση διαμοριακών
ενεργειών αλληλεπίδρασης τύπου OH..π
με την μέθοδο CASSCF/NEVPT2.**



3.1. Εισαγωγή

Σε μία βιβλιογραφική αναφορά ο Jeffrey¹¹³ ταξινομεί τους ασθενείς δεσμούς υδρογόνου ανάλογα με την ενέργεια τους, που είναι μικρότερη από 4 Kcal/mol. Ο ποσοτικός προσδιορισμός τέτοιων ασθενών δυνάμεων, μεταξύ μοριακών συστημάτων, είναι απαιτητικός τόσο πειραματικά όσο και θεωρητικά¹¹⁴. Η ακριβής διερεύνηση της μοριακής υπερεπιφάνειας δυναμικής ενέργειας απαιτεί την χρήση πολύ-οριζουσιακών κυματοσυναρτήσεων που συμπεριλαμβάνουν την ηλεκτρονική συσχέτιση. Οι πολύ-οριζουσιακές κυματοσυναρτήσεις αναπτύσσονται ως σειρά από ορίζουσες Slater, για την υπολογιστική μελέτη.

Η μελέτη π μοριακών συστημάτων με μονο-οριζουσιακές μεθόδους, όπως η Hartree-Fock, συνήθως είναι ανεπαρκείς διότι υπάρχουν περισσότερες της μίας ορίζουσες που συνεισφέρουν σε μεγάλο ποσοστό στην ολική ηλεκτρονική ενέργεια του μοριακού συστήματος. Κατά την διάρκεια της διαδικασίας SCF, ενός π μοριακού συστήματος, είναι επιθυμητό τα π μοριακά τροχιακά να αντιμετωπιστούν ισοδύναμα, γεγονός που προτρέπει στην χρήση πολύ-οριζουσιακών κυματοσυναρτήσεων. Η επέκταση των μονο-οριζουσιακών μεθόδων σε πολύ-οριζουσιακές γίνεται με την μέθοδο MCSCF (MultiConfigurational Self-Consistent Field). Η πιο γνωστή και εφαρμόσιμη παραλλαγή της θεωρίας MCSCF είναι η μέθοδος CASSCF¹¹⁵ (Complete Active Space Self-Consistent Field) η οποία καλείται και μέθοδος FORS¹¹⁶ (Full-Optimized Reaction Space).

Μια βελτιωμένη περιγραφή της πολύ-οριζουσιακής διαμοριακής δυναμικής επιφάνειας ενέργειας είναι δυνατόν να γίνει συμπεριλαμβάνοντας την δυναμική ηλεκτρονική συσχέτιση. Η μεγαλύτερη συνεισφορά της δυναμικής ηλεκτρονικής συσχέτισης ανακτάται με την εφαρμογή της θεωρία διαταραχών στην πολύ-οριζουσιακή κυματοσυναρτήση η οποία διατηρεί και την ιδιότητα της size-consistency¹¹⁷. Πολλές υπολογιστικές μέθοδοι έχουν αναπτυχθεί ώστε να επεκτείνουν της πολύ-οριζουσιακές κυματοσυναρτήσεις με την θεωρία διαταραχών δεύτερης τάξης. Η πιο γνωστή είναι η μέθοδος CASPT2^{118,119} (second-order perturbation theory based on a CASSCF reference wavefunction).

Σχετικά πρόσφατα μία νέα θεωρία αναπτύχθηκε η οποία επεκτείνει την πολύ-οριζουσιακές κυματοσυναρτήσεις με την θεωρία διαταραχών 2^{ης} τάξης και καλείται NEVPT2 (*n*-electron valence state perturbation theory)^{120,121,122}. Το μεγάλο πλεονέκτημα της μεθόδου είναι η διατήρηση της size-consistency του μοριακού συστήματος. Η μέθοδος στηρίζεται στην χαμιλτωνιανή του Dyal¹²³ και έχει



αναπτυχθεί στο πακέτο κβαντικών υπολογισμών Dalton¹²⁴.

Σε αυτό το κεφάλαιο εφαρμόζεται η μέθοδος NEVPT2 για την μελέτη της διαμοριακής ενέργειας αλληλεπίδρασης του ετεροδιμερούς $C_2H_4-H_2O$, ως μοντέλο μοριακού συστήματος για ασθενείς δεσμούς υδρογόνου του τύπου OH...π. Αναπτύσσεται μια υπολογιστική διαδικασία για τον προσδιορισμό του κατάλληλου ενεργού χώρου ώστε να εφαρμοστεί η μέθοδος CASSCF. Προκαταρκτικοί υπολογισμοί γίνονται με την μέθοδο CI, χρησιμοποιώντας μικρά σύνολα βάσεων, ώστε να εντοπιστούν τα ενεργά μοριακά τροχιακά (MOs) ενώ επιπλέον χρησιμοποιήθηκαν πακέτα οπτικοποίησης των μοριακών τροχιακών. Η διαδικασία εφαρμόζεται σε δύο διαφορετικούς ενεργούς χώρους και όλοι οι υπολογισμοί έγιναν με την μέθοδο NEVPT2, η οποία συγκλίνει σταθερά σε όλες τις περιπτώσεις και σε όλα τα μοντέλα που μελετήθηκαν.

Ο κύριος στόχος του κεφαλαίου αυτού είναι η ανάπτυξη μια μεθοδολογίας ώστε να γίνονται υπολογισμοί υψηλής ακρίβειας σε ετεροδιμερή. Τα ετεροδιμερή είναι συστήματα που αποτελούνται από δύο διαφορετικά αλληλεπιδρώντα μόρια. Η ιδιότητα της size-consistency των μεθόδων CASSCF και NEVPT2 επιτρέπει την αξιόπιστη εφαρμογή τέτοιων πολύ-οριζουσιακών υπολογισμών. Επίσης αναπτύσσεται μία διαδικασία που κάνει εφικτή την χρήση της μεθόδου CP των Boys και Bernardi για τον υπολογισμό του BSSE με την μέθοδο NEVPT2.

Η μεθοδολογία που αναπτύσσεται στο κεφάλαιο αυτό στοχεύει στην εφαρμογή της σε μεγαλύτερα και πολυπλοκότερα μοριακά συστήματα. Σε μία προηγούμενη μελέτη¹²⁵ διαπιστώθηκε δυσκολία ώστε να επιτευχθεί σύγκλιση με τη μέθοδο CASPT2 ενώ πρόσφατα μία μελέτη¹²⁶ με την μέθοδο CASPT2(12,12) στο διμερές της κυτοσύνης ολοκληρώθηκε επιτυχώς. Γενικά υπάρχει δυσκολία στον υπολογισμό CP-διορθωμένων ενεργειών σε ετεροδιμερή μοριακά συστήματα. Η μεθοδολογία που αναπτύσσεται σε αυτό το κεφάλαιο εφαρμόζεται επιτυχώς στην μελέτη του $C_6H_6-H_2O$ ως άλλο ένα παράδειγμα OH...π ασθενούς διαμοριακής αλληλεπίδρασης¹²⁷.

3.2. Δυναμική και μη-Δυναμική Ηλεκτρονική Συσχέτιση

Στη μέθοδο Hartree-Fock (HF) δίνεται η λύση της εξίσωσης Schrodinger θεωρώντας ότι η ηλεκτρονική αλληλεπίδραση αντικαθίσταται από ένα μέσο πεδίο χρησιμοποιώντας την διαδικασία SCF. Χρησιμοποιώντας ένα αρκετά μεγάλο σύνολο βάσης η μέθοδος HF είναι δυνατόν να υπολογίσει περίπου το 98% της συνολικής ηλεκτρονικής ενέργειας. Το υπολειπόμενο 2% είναι αρκετές φορές σημαντικό για την



περιγραφή διαφόρων χημικών φαινομένων. Η διαφορά της υπολογιζόμενης HF ηλεκτρονικής ενέργειας και της χαμηλότερης δυνατής (με θεωρητικά άπειρο σύνολο βάσης) ονομάζεται *ενέργεια ηλεκτρονικής συσχέτισης* (E_{cor})¹²⁸. Γενικά, η μέθοδος HF θεωρείται μια προσέγγιση της ολικής ηλεκτρονικής ενέργειας της βασικής κατάστασης, καθώς και μία πολύ καλή αρχή για την περαιτέρω δημιουργία μεθόδων υπολογισμού της ηλεκτρονικής συσχέτισης¹²⁹. Το ποσό της ενέργειας ηλεκτρονικής συσχέτισης που υπολογίζεται με την μέθοδο HF οφείλεται ως επί το πλείστον στις ηλεκτρονικές αλληλεπιδράσεις μικρής εμβέλειας (short-range). Υπάρχουν όμως αρκετές περιπτώσεις όπου η μέθοδος HF αποτυγχάνει να δώσει μια καλή φυσική περιγραφή του μοριακού συστήματος. Στις περιπτώσεις αυτές η ενέργεια ηλεκτρονικής συσχέτισης¹³⁰ πρέπει να θεωρηθεί ως το σύνολο που αποτελείται από μία μη-δυναμική συνεισφορά του συστήματος και μία δυναμική συνεισφορά^{131,132,133}. Η μη-δυναμική ενέργεια ηλεκτρονικής συσχέτισης αποδίδεται σε μεγάλης εμβέλειας (long-range) ηλεκτρονικές αλληλεπιδράσεις, που οφείλονται στην αλληλεπίδραση της βασικής κατάστασης HF με γειτονικές καταστάσεις.

Η μέθοδος HF συγκλίνει στην καλύτερη κυματοσυνάρτηση με το επιλεγμένο σύνολο βάσης των ατομικών τροχιακών. Για να βελτιωθούν τα ενεργειακά αποτελέσματα ως προς την E_{cor} είναι δυνατόν ως αρχική δοκιμαστική κυματοσυνάρτηση να οριστεί κάποια που να περιέχει περισσότερες οριζουσες Slater μοριακών τροχιακών. Οι μέθοδοι υπολογισμού της ενέργειας ηλεκτρονικής συσχέτισης E_{cor} χρησιμοποιούν την κυματοσυνάρτηση HF ως αρχή για περαιτέρω ενεργειακές βελτιώσεις. Άρα, μια αρχική και γενική πολύ-οριζουσιακή κυματοσυνάρτηση που μπορεί να χρησιμοποιηθεί έχει την μορφή:

$$\Psi = \alpha_0 \Phi_{\text{HF}} + \sum_{i=1} a_i \Phi_i \quad (3.1)$$

Οι μέθοδοι που χρησιμοποιούν την εξίσωση 3.1 για τον υπολογισμό της E_{cor} συνήθως διαφέρουν ως προς τον τρόπο που υπολογίζουν τους συντελεστές συμμετοχής των οριζουσών Φ_i για τους οποίους ισχύει η επόμενη συνθήκη κανονικοποίησης:

$$\sum_{i=0} a_i^2 = 1 \quad (3.2)$$

Οι τρεις κύριες μέθοδοι που χρησιμοποιούν ως αρχή τις εξισώσεις (3.1) και (3.2) για τον υπολογισμό της E_{cor} ονομάζονται: Αλληλεπίδραση σχηματισμών (Configuration Interaction - CI), Θεωρία διαταραχών πολλών σωμάτων (Many Body Perturbation Theory - MBPT) και Θεωρία Coupled Cluster (CC).



Στην περίπτωση που οι ορίζουσες Φ_i συμπεριλαμβάνουν διεγέρσεις όλων των εξωτερικών (valence) μοριακών τροχιακών, τότε η ενέργεια μη-δυναμικής ηλεκτρονικής συσχέτισης ορίζεται ως:

η διαφορά της πλήρους αλληλεπίδρασης σχηματισμών (*Full Configuration Interaction*) των οριζουσών Φ_i και της βασικής ορίζουσας Φ_{HF} των μοριακών τροχιακών.

Για την περιγραφή φαινομένων όπως η διάσπαση δεσμών, όπου η θεωρία HF αποτυγχάνει, η αξιόπιστη περιγραφή της μη-δυναμικής ηλεκτρονικής συσχέτισης είναι απαραίτητη¹³⁴.

3.3. Η μέθοδος CI (Configuration Interaction)

Η μέθοδος CI (Configuration Interaction) περιγράφεται από την κυματοσυνάρτηση Ψ_{CI} η οποία είναι ανάλογη της εξίσωσης (3.1) και είναι η εξής:

$$\Psi_{CI} = \alpha_0 \Phi_{HF} + \sum_S \alpha_S \Phi_S + \sum_D \alpha_D \Phi_D + \sum_T \alpha_T \Phi_T + \dots = \sum_{i=0} \alpha_i \Phi_i \quad (3.3)$$

Η κυματοσυνάρτηση Ψ_{CI} γράφεται ως γραμμικός συνδυασμός οριζουσών τύπου Slater και αφορά βελτιστοποίηση των συντελεστών α_i με στόχο την χαμηλότερη δυνατή ηλεκτρονική ενέργεια, με την προϋπόθεση να ισχύει η συνθήκη κανονικοποίησης των συντελεστών της εξίσωσης (3.2). Οι ορίζουσες Slater Φ_i για $i > 0$ που συμπεριλαμβάνονται στον γραμμικό συνδυασμό αποτελούνται από μονές (S), διπλές (D), τριπλές (T), κλπ διεγέρσεις των μοριακών τροχιακών της βασικής ορίζουσας Φ_{HF} .

Ο υπολογισμός της ενέργειας με την κυματοσυνάρτηση CI είναι αντίστοιχος με τον υπολογισμό της ενέργειας με την θεωρία HF. Πρέπει να επιλυθεί ένα σύνολο σαϊκουλικών εξισώσεων το οποίο είναι ισοδύναμο με την διαγωνοποίηση ενός πίνακα ο οποίος λέγεται πίνακας CI¹²⁸. Αν όλες οι ορίζουσες συμπεριληφθούν στην εξίσωση (3.3) έχουμε μία πλήρη κυματοσυνάρτηση (Full CI) η οποία στο επιλεγμένο σύνολο βάσης αποδίδει το 100% της ενέργειας ηλεκτρονικής συσχέτισης. Αυτό είναι εφικτό για πάρα πολύ μικρά μόρια και μόνο. Στους πρακτικούς υπολογισμούς η εξίσωση (3.3) περιορίζεται στους δύο έως τέσσερις πρώτους όρους, συμπεριλαμβάνοντας μόνο τις ορίζουσες με μονές, διπλές, τριπλές και τετραπλές διεγέρσεις ανάλογα με το σύνολο βάσης και το σύστημα προς μελέτη. Ο περιορισμός αυτός αποδίδει το 80% έως 90% της ενέργειας ηλεκτρονικής συσχέτισης.



3.4. Η μέθοδος MCSCF

Αν και οι πρώτες αναφορές στην μέθοδο MCSCF έγιναν την δεκαετία του 1930^{135,136}, η πρώτη πρακτική εφαρμογή της έγινε στα μέσα του 1960 με το μοντέλο OVC (Optimized Valence Configuration) των Wahl, Das και Gilbert^{137,138}.

Η SCF (Self Consistent Field) μέθοδος των πολλαπλών σχηματισμών (Multi Configuration Self Consistent Field - MCSCF) περιγράφεται από την κυματοσυνάρτηση (3.3) της CI με την διαφορά ότι εκτός από τους συντελεστές α_i βελτιστοποιούνται και τα μοριακά τροχιακά που περιέχονται στις ορίζουσες¹³³. Η διαδικασία της βελτιστοποίησης είναι επαναληπτική όπως και η μέθοδος SCF και στην περίπτωση που η εξίσωση (3.3) έχει έναν μόνο όρο τότε η μέθοδος MCSCF ταυτίζεται με την HF. Η συνθήκη κανονικοποίησης των συντελεστών της εξίσωσης (3.2) πρέπει να ισχύει και στην κυματοσυνάρτηση MCSCF.

Στην διαδικασία SCF, μόνο τα κατειλημμένα (occupied) μοριακά τροχιακά συνεισφέρουν στην ολική ηλεκτρονική ενέργεια. Εφόσον η φύση του προβλήματος (διάσπαση δεσμών, διεγερμένες καταστάσεις, κλπ) επιβάλλει και την συμβολή μη-κατειλημμένων μοριακών τροχιακών στην κυματοσυνάρτηση είναι απαραίτητη η χρήση της μεθόδου MCSCF. Έστω ότι ένα μοριακό σύστημα αποτελείται από N ηλεκτρόνια και M μοριακά τροχιακά. Τα K είναι τα κατειλημμένα μοριακά τροχιακά και τα $M-K$ είναι τα μη-κατειλημμένα. Εφόσον χρειάζεται το μοριακό τροχιακό $K+1$ να συνεισφέρει στην ολική ηλεκτρονική ενέργεια θα χρησιμοποιήσουμε την εξής MCSCF κυματοσυνάρτηση:

$$\Psi_{MCSCF} = \alpha_0 |\dots \phi_{K-1}^2 \phi_K^2 \rangle + \alpha_1 |\dots \phi_{K-1}^2 \phi_{K+1}^2 \rangle \quad (3.4)$$

η οποία αποτελείται μόνο από δύο σχηματισμούς (ορίζουσες) και καλείται TCSCF (Two Configurations SCF). Ο όρος σχηματισμός (Configuration State Function - CSF) χρησιμοποιείται ως γενικότερος της ορίζουσας τύπου Slater καθότι ένας σχηματισμός μπορεί να είναι γραμμικός συνδυασμός πολλών ορίζουσών σε μία κυματοσυνάρτηση MCSCF. Ένας σχηματισμός διατηρεί το ολικό spin S του συστήματος για αυτό μπορεί να είναι γραμμικός συνδυασμός δύο ορίζουσών μονοηλεκτρονικών διεγέρσεων σε ένα μοριακό σύστημα με ολικό spin $S=1$ (singlet). Με αυτό τον τρόπο οι σχηματισμοί που συμμετέχουν σε μία κυματοσυνάρτηση MCSCF στην οποία ο ενεργός χώρος αποτελείται από N ηλεκτρόνια και K μοριακά τροχιακά δίνεται από το επόμενο τύπο των Weyl-Robinson¹³⁹:



$$CSFs = \frac{2S+1}{K+1} \binom{K+1}{K-1/2N-S} \binom{K+1}{1/2N-S} \quad (3.5)$$

όπου S είναι το ολικό spin.

Ο αριθμός των σχηματισμών μπορεί να είναι υπερβολικά μεγάλος σύμφωνα με την εξίσωση (3.5) καθώς εξαρτάται από τα ηλεκτρόνια και τα μοριακά τροχιακά που θα συμπεριληφθούν στην κυματοσυνάρτηση MCSCF.

Ένα μοριακό τροχιακό μπορεί να εμφανίζεται σε περισσότερους από έναν σχηματισμούς της κυματοσυνάρτησης MCSCF και να καταλαμβάνεται από κανένα, ένα ή δύο ηλεκτρόνια. Τα μοριακά τροχιακά που προκύπτουν από την MCSCF δεν έχουν ενεργειακές ιδιοτιμές αλλά αριθμούς κατάληψης (occupation numbers) στο κλειστό διάστημα $[0..2]$.

3.4.1. Υπολογισμός της MCSCF ενέργειας

Ο υπολογισμός της MCSCF ηλεκτρονικής ενέργειας γίνεται ως εξής:

$$\langle E \rangle = \langle \Psi_{MCSCF} | H | \Psi_{MCSCF} \rangle = \sum_{K,L} \alpha_K \alpha_L \langle \psi_{MCSCF}^K | H | \psi_{MCSCF}^L \rangle$$

$$\frac{\partial \langle E \rangle}{\partial \alpha_K} = 0, \dots$$

Η επίλυση των παραπάνω εξισώσεων ιδιοτιμών δίνει:

- Τις MCSCF ενέργειες E_N για κάθε ηλεκτρονική κατάσταση.
- Τους συντελεστές $\alpha_{J,N}$ της CI για κάθε κατάσταση N .

Η υπολογιστική εύρεση της ολικής MCSCF ηλεκτρονικής ενέργειας γίνεται με επαναληπτική διαδικασία όπου:

- Για κάθε σύνολο συντελεστών α_K , υπολογίζονται οι συντελεστές C_μ των μοριακών τροχιακών που είναι LCAO (micro-iterations).
- Για ένα δοσμένο σύνολο C_μ επιλύεται η πλήρης CI στον ενεργό χώρο για να βρεθούν οι νέοι συντελεστές α_K (macro-iterations).
- Η διαδικασία επαναλαμβάνεται μέχρι να επέλθει αυτοσυνέπεια (self-consistency) στο σύστημα.



3.4.2. Σύγκλιση της MCSCF

Σε αντίθεση με την σύγκλιση της διαδικασίας SCF η σύγκλιση της MCSCF είναι πιθανόν να οδηγήσει σε ενεργειακό αποτέλεσμα το οποίο δεν είναι ελάχιστο. Κατά την διάρκεια της σύγκλισης είναι δυνατόν τα κριτήρια σύγκλισης να ικανοποιούνται αλλά η MCSCF κυματοσυνάρτηση να μην είναι βελτιστοποιημένη (συντελεστές και μοριακά τροχιακά).

Η καλύτερη δυνατή σύγκλιση της MCSCF αντιμετωπίζεται με μεθόδους βελτιστοποίησης δεύτερης τάξης, χρησιμοποιώντας την πληροφορία του θετικά ορισμένου εσσιανού πίνακα. Αυτό επιτυγχάνεται με την βελτιστοποίηση των συντελεστών και των τροχιακών χρησιμοποιώντας την μέθοδο Newton-Raphson. Η μέθοδος αυτή για την σύγκλιση της MCSCF εφαρμόστηκε για πρώτη φορά το 1978 από τους Dalgaard και Jorgensen¹⁴⁰ και χρησιμοποιείται στους πρακτικούς υπολογισμούς μέχρι σήμερα.

3.4.3. Απόδοση της MCSCF

Η μέθοδος MCSCF συνήθως χρησιμοποιείται για να δώσει περισσότερη ακρίβεια στην ενέργεια ηλεκτρονικής συσχέτισης. Η ακρίβεια αυτή εξαρτάται από το πλήθος των σχηματισμών που συμπεριλαμβάνονται στην κυματοσυνάρτηση και ο στόχος είναι η καλύτερη ποιοτική περιγραφή του πολύ-ηλεκτρονιακού συστήματος. Προσθέτοντας περισσότερους σχηματισμούς CSFs στην κυματοσυνάρτηση η MCSCF συγκλίνει σε χαμηλότερη ολική ηλεκτρονική ενέργεια. Η επιπρόσθετη ηλεκτρονική ενέργεια που εισάγεται είναι η ενέργεια μη-δυναμικής ηλεκτρονικής συσχέτισης. Αυτό οφείλεται στο γεγονός ότι η MCSCF κυματοσυνάρτηση είναι πιο ευέλικτη καθώς περιέχει ορίζουσες με τροχιακά τα οποία είναι μονο-ηλεκτρονιακά κατελιημμένα. Η υπόλοιπη ποσότητα της ενέργειας ηλεκτρονικής συσχέτισης είναι δυναμικού τύπου και συνήθως δεν είναι εφικτός ο διαχωρισμός των δύο αυτών τύπων. Η μέθοδος MCSCF χρησιμοποιείται για την δημιουργία ποιοτικά καλύτερων κυματοσυναρτήσεων με στόχο την ανάκτηση της ενέργειας μη-δυναμικής ηλεκτρονικής συσχέτισης.

Αν στην MCSCF συμπεριληφθούν όλοι οι όροι της εξίσωσης (3.3) τότε έχουμε μια πλήρη κυματοσυνάρτηση CI, όπου όλα τα μοριακά τροχιακά συμπεριλαμβάνονται και δεν χρειάζεται να βελτιστοποιηθούν (είναι βελτιστοποιημένα από την HF). Όμως αυτό στους πρακτικούς υπολογισμούς δεν είναι δυνατόν να εφαρμοστεί και σε αυτό το σημείο συμβάλει η μέθοδος MCSCF.



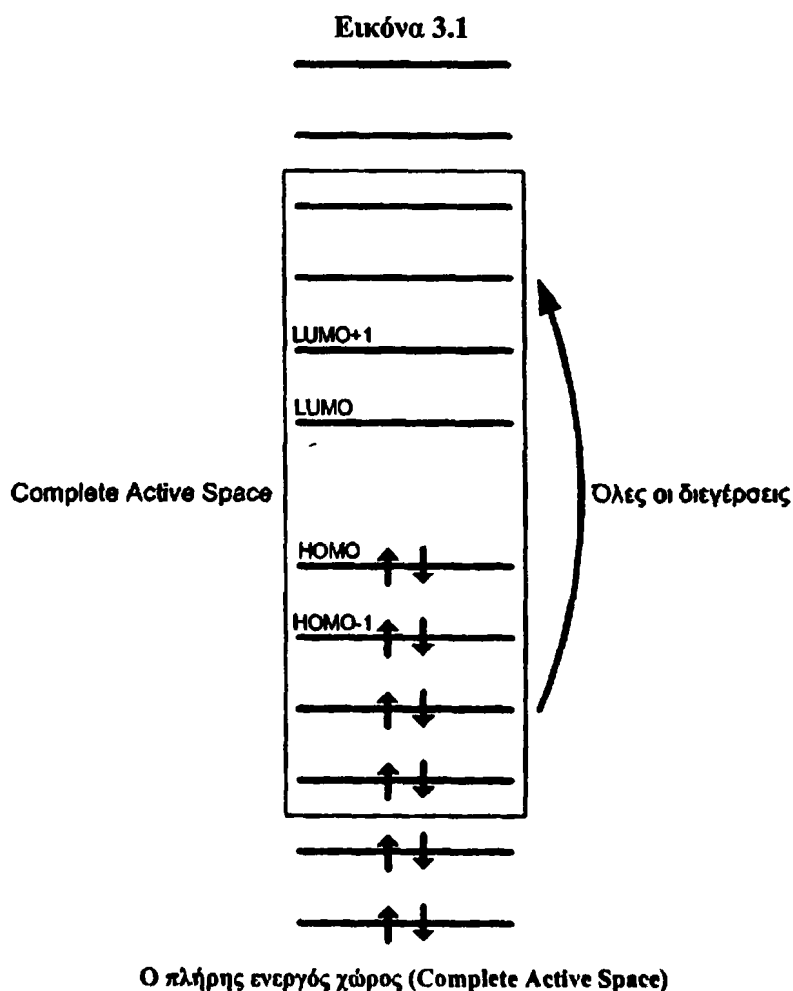
3.4.4. Περιορίζοντας το μέγεθος της MCSCF

Το κυρίως πρόβλημα της μεθόδου MCSCF ανάγεται στην επιλογή των σχηματισμών CSFs που θα συμμετέχουν στην κυματοσυνάρτηση. Θεωρώντας ότι οι μοριακές αλληλεπιδράσεις εκφράζονται από μεταβολές στα εξωτερικά μοριακά τροχιακά, ο υπολογισμός της ηλεκτρονικής ενέργειας συσχέτισης μπορεί να επιτευχθεί με κυματοσυναρτήσεις που συνδέονται με τα εξωτερικά ηλεκτρόνια. Η θεώρηση αυτή οδηγεί σε περιορισμένη επέκταση των σχηματισμών που συμμετέχουν στην κυματοσυνάρτηση MCSCF. Όμως η επέκταση αυτή πραγματοποιείται συμπεριλαμβάνοντας όλους τους δυνατούς συνδυασμούς των εξωτερικών μοριακών τροχιακών, εκφράζοντας τους όρους μιας πλήρους CI (Full Configuration Interaction) με τα επιλεγμένα μοριακά τροχιακά. Μία απευθείας διαδικασία για να ποσοτικοποιηθεί αυτή η μέθοδος προτάθηκε από τους Ruedenberg και Roos με την ονομασία Full Optimized Reaction Space (FORS) ή ισοδύναμα Complete Active Space (CASSCF) όπως αναφέρθηκε στην εισαγωγή. Η μέθοδος CASSCF δίνει την μέγιστη ενέργεια ηλεκτρονικής συσχέτισης των εξωτερικών (valence) μοριακών τροχιακών που συμμετέχουν στην κυματοσυνάρτηση.



3.5. Η μέθοδος CASSCF(n,m)

Στην μέθοδο CASSCF η επιλογή των σχηματισμών CSFs που θα συμπεριληφθούν στην επέκταση της εξίσωσης (3.4), ανάγεται στον διαχωρισμό των μοριακών τροχιακών σε ενεργά (active) και ανενεργά (inactive). Ο ενεργός χώρος (Active Space) στον οποίο εφαρμόζεται η κυματοσυνάρτηση MCSCF αποτελείται μόνο από τα ενεργά μοριακά τροχιακά. Τα ενεργά μοριακά τροχιακά επιλέγονται μεταξύ των υψηλότερα κατειλημμένων (HOMO) και χαμηλότερα μη-κατειλημμένων (LUMO). Τα ανενεργά μοριακά τροχιακά είναι συνήθως όλα τα υπόλοιπα με καταληψιμότητα από δύο ή κανένα ηλεκτρόνια. Αυτό φαίνεται στην εικόνα 3.1.



Στον ενεργό χώρο όλοι οι δυνατοί σχηματισμοί CSFs ταυτίζονται με τους σχηματισμούς που παράγει μια πλήρης κυματοσυνάρτηση CI. Αν πάρουμε για παράδειγμα την εικόνα 3.1 όπου ο ενεργός χώρος αποτελείται από 8 μοριακά τροχιακά και 8 ηλεκτρόνια σύμφωνα με την εξίσωση (3.5) έχουμε 1764 CSFs,

θεωρώντας το ολικό spin $S=1$ (singlet). Με αυτούς τους 1764 όρους δημιουργείται η πολύ-οριζουσιακή κυματοσυνάρτηση MCSCF η οποία για την συγκεκριμένη περίπτωση συμβολίζεται ως CASSCF(8,8). Γενικότερα χρησιμοποιείται ο συμβολισμός CASSCF(n,m), όπου n είναι ο αριθμός των ηλεκτρονίων και m ο αριθμός των ενεργών μοριακών τροχιακών.

3.5.1. Οι υπολογισμοί με την μέθοδο CASSCF

- Στην περίπτωση της CASSCF γίνεται βελτιστοποίηση των συντελεστών της CI και των μοριακών τροχιακών, όπως περιγράφηκε στην ενότητα 3.4.1.
- Η βελτιστοποίηση των συντελεστών των μοριακών τροχιακών είναι ανάλογη αλλά πολυπλοκότερη από την HF, διότι χρειάζεται να βελτιστοποιηθούν μοριακά τροχιακά που ανήκουν στον core, active και inactive χώρο τα οποία αναμιγνύονται. Για τον λόγο αυτό η CASSCF έχει και μεγαλύτερο υπολογιστικό κόστος σε σχέση με την MCSCF.
- Η βελτιστοποίηση των μοριακών τροχιακών στην περίπτωση της CASSCF είναι ακριβώς ανάλογη με την γεωμετρική βελτιστοποίηση όπου η παράγωγος των μοριακών τροχιακών ως προς την θέση τους (orbital rotation) αναπτύσσεται σε μία σειρά Taylor.

3.5.2. Η εφαρμογή της μεθόδου CASSCF

Στην συνέχεια δίνονται μερικά βασικά χαρακτηριστικά για την εφαρμογή της μεθόδου CASSCF, δεδομένου ότι ο αριθμός και ο τύπος των μοριακών τροχιακών που περιλαμβάνονται στον ενεργό χώρο εξαρτάται από την φύση του προβλήματος και λαμβάνεται υπ' όψη το υπολογιστικό κόστος.

- Συνήθως ο ενεργός χώρος αποτελείται από όλα τα μοριακά τροχιακά τα οποία αναμένεται να συνεισφέρουν στην ενέργεια ηλεκτρονικής συσχέτισης. Η εισαγωγή ενός μεγάλου αριθμού μη-κατειλημμένων μοριακών τροχιακών στην μέθοδο CASSCF οδηγεί στην εκχώρηση ποσότητας της δυναμικής ενέργειας συσχέτισης στην μη-δυναμική^{141,142,143,144,145}.
- Η ανάπτυξη των όρων της πλήρης CI που αντιστοιχεί στον ενεργό χώρο με n ηλεκτρόνια σε m μοριακά τροχιακά είναι συνήθως πολύ μεγάλος για τους πρακτικούς υπολογισμούς. Αυτό οδηγεί στον περιορισμό των n και m σε τιμές



γύρω από το 18. Η επιλογή των «απαραιτήτων» μοριακών τροχιακών που περιγράφουν το φυσικό πρόβλημα είναι συνήθως αναγκαία.

- Η εξέταση των συντελεστών a_i της αναπτυγμένης κυματοσυνάρτησης μας δίνει την ποσοτική πληροφορία της συμμετοχής των επιλεγμένων σχηματισμών CSFs μόνο σε κυματοσυναρτήσεις της μορφής CI. Στις κυματοσυναρτήσεις MCSCF και CASSCF, λόγω της διαδικασίας βελτιστοποίησης (orbital rotation), οι συντελεστές είναι σε κάποιο βαθμό αυθαίρετοι οπότε δεν πρέπει να λαμβάνονται υπ' όψη για να οριστεί ο πολύ-οριζουσιακός χαρακτήρας του προβλήματος.
- Η μέθοδος CASSCF αποδίδει την μέγιστη ενέργεια ηλεκτρονικής συσχέτισης μόνο του ενεργού χώρου και όχι του ανενεργού¹⁴⁶. Αυτό επιτυγχάνεται αν όλα τα εξωτερικά ηλεκτρόνια συμπεριληφθούν στον ενεργό χώρο, το οποίο είναι δυνατόν για σχετικά μικρά μοριακά συστήματα.
- Ο πολύ-οριζουσιακός χαρακτήρας του φυσικού προβλήματος επιβεβαιώνεται από την παρουσία αριθμών κατάληψης στα μοριακά τροχιακά μεταξύ 0.1 και 1.9.

Για την εφαρμογή της μεθόδου CASSCF χρειάζεται ιδιαίτερη προσοχή στην επιλογή του ενεργού χώρου και πιο συγκεκριμένα στην επιλογή των ενεργών μοριακών τροχιακών.

3.5.3. Σχεδιάζοντας τον κατάλληλο ενεργό χώρο

Ο σχεδιασμός του κατάλληλου ενεργού χώρου (Active Space) ανάγεται στην επιλογή των μοριακών τροχιακών που θα περιλαμβάνονται σε αυτόν. Υπάρχουν μερικοί εμπειρικοί κανόνες οι οποίοι μπορούν να εφαρμοστούν για την επιλογή των κατάλληλων μοριακών τροχιακών. Οι κανόνες αυτοί δεν μπορούν να εφαρμοστούν σε όλες τις περιπτώσεις και έχουν προκύψει από εμπειρικούς υπολογισμούς.

1. Για κάθε κατειλημμένο μοριακό τροχιακό αντιστοιχεί περίπου ένα μη-κατειλημμένο μοριακό τροχιακό. Δηλαδή ο ενεργός χώρος αποτελείται από περίπου ίσο αριθμό κατειλημμένων και μη-κατειλημμένων μοριακών τροχιακών (εικόνα 3.1).
2. Οι ενέργειες των κανονικών μοριακών από έναν RHF υπολογισμό μπορούν να χρησιμοποιηθούν ως κριτήριο επιλογής. Συνήθως τα HOMO και LUMO είναι τα πιο σημαντικά τροχιακά που συμπεριλαμβάνονται στον ενεργό χώρο. Αυτό



επιβεβαιώνεται και από την θεωρία διαταραχών 2^{ης} τάξης όπου μικρή διαφορά στα μοριακά τροχιακά έχει μεγάλη συνεισφορά στην ενέργεια ηλεκτρονικής συσχέτισης. Η θεώρηση αυτή όμως χρειάζεται ιδιαίτερη προσοχή στην περίπτωση που το φυσικό πρόβλημα έχει πολύ-οριζουσιακό χαρακτήρα και τα RHF μοριακά τροχιακά μόνο από την βασική ορίζουσα δεν μπορούν να περιγράψουν την σωστή κυματοσυνάρτηση - διότι δεν είναι ακόμα γνωστή.

3. Ένας αρκετά αποδοτικός τρόπος στηρίζεται στην χρήση των φυσικών τροχιακών (natural orbitals)¹⁴⁷. Τα φυσικά τροχιακά εξάγονται από την διαγωνοποίηση του πίνακα ηλεκτρονικής πυκνότητας και οι ιδιοτιμές που προκύπτουν είναι οι αριθμοί κατάληψης των φυσικών τροχιακών (Natural Orbitals Occupation Numbers - NOON). Το άθροισμα των αριθμών κατάληψης είναι ίσο με το σύνολο των ηλεκτρονίων του συστήματος. Φυσικά τροχιακά με αριθμούς κατάληψης αρκετά διαφορετικούς από 0 και 2 είναι συνήθως αυτά που πρέπει να συμπεριληφθούν στον ενεργό χώρο. Μία κυματοσυνάρτηση RHF έχει αριθμούς κατάληψης μόνο 0 ή 2. Αν προστεθεί ηλεκτρονική συσχέτιση στην κυματοσυνάρτηση RHF τότε προκύπτουν αριθμοί κατάληψης διαφορετικοί από 0 και 2. Αυτό μπορεί να γίνει εκτελώντας προκαταρκτικούς υπολογισμούς MP2 ή CISD πριν την εφαρμογή της CASSCF.
4. Άλλη μία μεθοδολογία¹⁴⁸ στηρίζεται στους αριθμούς κατάληψης των φυσικών τροχιακών χρησιμοποιώντας προκαταρκτικούς υπολογισμούς CI με όλους τους σχηματισμούς που περιλαμβάνουν μόνο τις διπλές διεγέρσεις των μοριακών τροχιακών. Η επιλογή των μοριακών τροχιακών που θα συμμετέχουν στον ενεργό χώρο της CASSCF γίνεται με κριτήριο τους αριθμούς κατάληψης των φυσικών τροχιακών που προέκυψαν από την CI και πρέπει να είναι από 0.01 έως 1.99.

Όλοι οι παραπάνω κανόνες, στις περισσότερες περιπτώσεις, λειτουργούν συνδυαστικά και όπως έχει αποδειχτεί από πρακτικούς υπολογισμούς είναι απαραίτητη η χρήση προγράμματος οπτικοποίησης¹⁴⁹ των μοριακών τροχιακών που θα συμπεριληφθούν στην κυματοσυνάρτηση CASSCF.



3.6. Θεωρία διαταραχών 2^{ης} τάξης στην CASSCF

Όπως προαναφέρθηκε η μέθοδος CASSCF αν και είναι αρκετά ευέλικτη και αποδοτική δίνει μόνο ένα μέρος της ενέργειας ηλεκτρονικής συσχέτισης¹⁵⁰. Υπάρχει η δυνατότητα να υπολογιστεί μεγάλο μέρος της υπόλοιπης ηλεκτρονικής ενέργειας συσχέτισης χρησιμοποιώντας την θεωρία διαταραχών 2^{ης} τάξης (second order perturbation theory) στις κυματοσυναρτήσεις CASSCF. Η δυναμική ενέργεια ηλεκτρονικής συσχέτισης έχει αποδοθεί με μεγάλη επιτυχία^{151,152,153,154,155} χρησιμοποιώντας θεωρία διαταραχών δεύτερης τάξης σε πολύ-οριζουσιακές κυματοσυναρτήσεις όπως η CASSCF.

Οι πιο διαδεδομένες μέθοδοι εφαρμογής της θεωρίας διαταραχών δεύτερης τάξης σε κυματοσυναρτήσεις CASSCF είναι οι: CASPT2^{156,157}, MRMP/MCQDPT^{158,159,160,161} και NEVPT2, οι οποίες περιγράφονται στην συνέχεια. Παρόλο η ολικές ηλεκτρονικές ενέργειες είναι διαφορετικές με την εφαρμογή κάθε μεθόδου, οι σχετικές ενεργειακές διαφορές είναι σχεδόν παρόμοιες. Η εφαρμογή των μεθόδων αυτών επεκτείνεται όλο και περισσότερο σε μεγαλύτερα και πολυπλοκότερα μοριακά συστήματα όσο η υπολογιστική ισχύς αυξάνει.

3.6.1. CASPT2

Η πιο διαδεδομένη μέθοδος που χρησιμοποιείται ευρέως¹⁶² στους μοριακούς κβαντικούς υπολογισμούς είναι η μέθοδος CASPT2 η οποία χρησιμοποιεί ως αναφορά μία κυματοσυνάρτηση CASSCF. Η μέθοδος αναπτύχθηκε το 1990 από τον Roos, και χρησιμοποιεί ως μηδενικής τάξης κυματοσυνάρτηση την CASSCF για την ανάπτυξη των όρων της διαταραχής 2^{ης} τάξης. Στην μέθοδο CASPT2 ο χώρος των αλληλεπιδράσεων συμπεριλαμβάνεται σε μία πρώτης τάξης κυματοσυνάρτηση. Ο χαμιλτωνιανός τελεστής μηδενικής τάξης δημιουργείται από έναν μονο-ηλεκτρονιακό τελεστή Fock ο οποίος ισοδυναμεί με έναν Hartree-Fock Moller-Plesset τελεστή στην περίπτωση που δεν υπάρχουν ενεργά μοριακά τροχιακά.

Η πρώτης τάξης κυματοσυνάρτηση προκύπτει με μια επαναληπτική διαδικασία όπου λύνονται μεγάλα σύνολα γραμμικών εξισώσεων. Το μέγεθος του συστήματος των εξισώσεων είναι της τάξης του $n^2 \cdot m^2 / 2$, όπου n είναι το άθροισμα των ενεργών και ανενεργών μοριακών τροχιακών και m είναι το άθροισμα των ενεργών και δευτερευόντων (virtual) μοριακών τροχιακών. Για να εφαρμοστεί η CASPT2 πρέπει ο να έχει προηγηθεί αρχικός CASSCF υπολογισμός με όσο το δυνατόν καλύτερο



επιλεγμένο ενεργό χώρο. Ένα άλλο χαρακτηριστικό της μεθόδου CASPT2 είναι η «συνέπεια ως προς το μέγεθος» του μοριακού συστήματος (size-consistent).

Επέκταση της μεθόδου σε διαταραχές 3^{ης} τάξης (CASPT3), έδωσε^{163,164} ακόμα καλύτερα αποτελέσματα (σε μικρά μοριακά συστήματα) σε σύγκριση με τους υπολογισμούς 2^{ης} τάξης, αλλά με πολύ μεγαλύτερο υπολογιστικό κόστος.

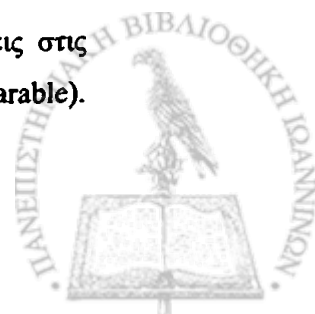
3.6.2. MRMP/MCQDPT

Η μέθοδος MRMP (Multi Reference Moller-Plesset) και MCQDPT (Multi-Configuration Quasi-Degenerate Perturbation Theory) αναπτύχθηκε από τους Hiraο το 1992 και Nakano το 1993 αντίστοιχα. Στην μέθοδο αυτή ο μονό-ηλεκτρονιακός τελεστής είναι ένας γενικευμένος τελεστής Fock που αντιστοιχεί στην κυματοσυνάρτηση αναφοράς MCSCF. Ένα μειονέκτημα είναι ο τελεστής αυτός δεν έχει φυσική σημασία όπως ο τελεστής Fock σε σύστημα κλειστών στοιβάδων (closed-shell). Τα κανονικά μοριακά τροχιακά μετασχηματίζονται σε φυσικά μοριακά τροχιακά πριν την ανάπτυξη των όρων της διαταραχής, οπότε η μονο-ηλεκτρονιακές αλληλεπιδράσεις θεωρούνται ως προς την κυματοσυνάρτηση MCSCF και μπορούν να αναπτυχθούν και στον χώρο των μη-κατειλημμένων μοριακών τροχιακών. Η μέθοδος MCQDPT είναι μία επέκταση της γενικής μεθόδου MRMP και μπορεί να εφαρμοστεί για την μελέτη διεγερμένων καταστάσεων όπου αποδίδει ταυτόχρονη λύση για πολλές καταστάσεις. Στις εφαρμογές της παρούσας διατριβής εφαρμόζεται για μία μόνο κατάσταση η μέθοδος MRMP. Η μέθοδος όπως αναφέρεται στην βιβλιογραφία είναι size-consistent υπό κάποιες προϋποθέσεις που δεν την κάνουν γενική στην χρήση της¹⁶⁵. Στην πράξη αποδείχτηκε να δίνει αρκετά έγκυρα αποτελέσματα και να αναδεικνύεται ως μια αξιόπιστη μέθοδο για τον υπολογισμό της ενέργειας δυναμικής συσχέτισης.

3.6.3. Προβλήματα των CASPT2 και MRMP/MCQDPT

Οι μέθοδοι CASPT2 και MCQDPT αντιμετωπίζουν το πρόβλημα της *εισβολής καταστάσεων* (intruder states) κατά την διαδικασία του ενεργειακού υπολογισμού της διαταραχής. Οι intruder states αποτελούν μία ιδιαίτερη περίπτωση που εμφανίζεται στους υπολογισμούς διαταραχών και όροι ανώτερης τάξης έχουν ενέργεια συγκρίσιμη με τον όρο μηδενικής τάξης.

Ένα άλλο πρόβλημα των μεθόδων CASPT2 και MRMP αφορά περιπτώσεις στις οποίες οι ηλεκτρονικές ενέργειες δεν είναι αυστηρά διαχωρίσιμες (strictly separable).



Αυτό σημαίνει ότι η ενέργεια μίας κατάστασης ενός μοριακού συστήματος A+B δεν ισούται με το ενεργειακό άθροισμα των αντίστοιχων καταστάσεων των A και B όταν η αλληλεπίδραση παύει να υπάρχει.

3.6.4. NEVPT2

Η μέθοδος NEVPT2 (n-electron valence space second-order perturbation theory) αναπτύχθηκε το 2001 για τον υπολογισμό της ενέργειας δυναμικής συσχέτισης σε κυματοσυναρτήσεις CASSCF χρησιμοποιώντας όπως και οι δύο προηγούμενες θεωρία διαταραχών 2^{ης} τάξης. Είναι παρόμοια με την μέθοδο CASPT2 με την διαφορά ότι στηρίζεται σε μία δι-ηλεκτρονιακή χαμιλτονιανή (Dyall Hamiltonian). Η μέθοδος NEVPT2 δεν εισάγει τα δύο προβλήματα που αναφέρθηκαν στην προηγούμενη ενότητα και για όλες τις περιπτώσεις εγγυάται την size-consistency του μοριακού συστήματος.

Η μέθοδος NEVPT2 στηρίζεται σε μία CAS-CI κυματοσυνάρτηση μηδενικής τάξης $\Psi_m^{(0)}$ όπου οι δι-ηλεκτρονιακές αλληλεπιδράσεις συμπεριλαμβάνονται και εκτός του ενεργού χώρου (CAS). Ο μη-ενεργός χώρος αποτελείται από τα εσωτερικά (core) και μη-κατειλημμένα ηλεκτρόνια (virtual) ενώ ο χώρος CAS από τα ενεργά μοριακά τροχιακά.

Έστω ότι έχουμε c μη-ενεργά και ν ενεργά ηλεκτρόνια. Τότε η κυματοσυνάρτηση μηδενικής τάξης γράφεται ως $|\Psi_m^{(0)}\rangle = |\Phi_c \Psi_m^v\rangle$, όπου Φ_c είναι το αντισυμμετρικό γινόμενο των ατομικών τροχιακών του μη-ενεργού χώρου και Ψ_m^v είναι η κυματοσυνάρτηση του ενεργού χώρου.

Εφόσον η μέθοδος αφορά διαταραχή 2^{ης} τάξης, οι κυματοσυναρτήσεις της διαταραχής είναι της μορφής $\Psi_m^{(k)}$, όπου $-2 \leq k \leq 2$. Αυτό σημαίνει ότι k ηλεκτρόνια μεταφέρονται από τον μη-ενεργό χώρο στον ενεργό χώρο. Πιο συγκεκριμένα οι όροι της διαταραχής είναι της μορφής $|\Psi_{l,\mu}^{(k)}\rangle = |\Phi_l^{-k} \Psi_\mu^{v+k}\rangle$ όπου το Φ_l^{-k} αποτελείται από c-k ηλεκτρόνια, οι κυματοσυναρτήσεις Ψ_μ^{v+k} από v+k ηλεκτρόνια, ο δείκτης l συμβολίζει τα μη ενεργά μοριακά τροχιακά και ο δείκτης μ απαριθμεί τις διαφορετικές κυματοσυναρτήσεις.

Ο υπολογισμός της $|\Psi_{l,\mu}^{(k)}\rangle = |\Phi_l^{-k} \Psi_\mu^{v+k}\rangle$ γίνεται με διαγωνοποίηση του χαμιλτωνιανού τελεστή στον χώρο S_l^k που περιλαμβάνει όλες τις οριζουσες με το ίδιο μη ενεργό



Φ_i^k και όλες τις πιθανές κυματοσυναρτήσεις Ψ_i^k στον ενεργό χώρο I. Η προσέγγιση αυτή ονομάζεται “totally uncontracted” διότι εκτείνεται σε όλο τον χώρο S_i^k και γενικά δεν είναι εφαρμόσιμη λόγω του πολύ υψηλού υπολογιστικού κόστους.

Μία άλλη προσέγγιση μπορεί να επιτευχθεί επιλέγοντας μία συνάρτηση από κάθε χώρο S_i^k . Ένας τελεστής της διαταραχής χρησιμοποιείται για να παράγει μία συνάρτηση από κάθε χώρο. Η προσέγγιση αυτή η οποία έχει το λιγότερο υπολογιστικό κόστος ονομάζεται Strongly Contracted (SC).

Η τελευταία προσέγγιση συνιστά την εύρεση ενός υποχώρου του S_i^k ο οποίος να αλληλεπιδρά με την κυματοσυνάρτηση $|\Psi_m^{(0)}\rangle$, μεθοδολογία που είναι συγγενείς με την μέθοδο CASPT2. Η προσέγγιση αυτή ονομάζεται Partially Contracted (PC) και έχει μεγαλύτερη διάσταση από την SC (είναι μονοδιάστατη σε κάθε υποχώρο). Στον υποχώρο που δημιουργείται από την PC αντιστοιχούν περισσότερες συναρτήσεις από την μία που έχει η SC σε κάθε χώρο S_i^k . Η προσέγγιση PC έχει υπολογιστικό κόστος της τάξης του N^3 , όπου N είναι ο αριθμός των ηλεκτρονίων του ενεργού χώρου. Αν και η μέθοδος SC παρέχει μικρότερη ευελιξία στους υπολογισμούς τα αποτελέσματα είναι σε εξίσου καλά με την προσέγγιση PC. Οι δύο προσεγγίσεις διαφέρουν στον αριθμό των διαταραγμένων συναρτήσεων που εμφανίζονται στην σειρά της διαταραχής. Παρόλο που τα αποτελέσματα των δύο προσεγγίσεων είναι γενικά κοντά, η προσέγγιση PC είναι πιο ακριβής διότι έχει πιο εκτεταμένο χώρο συναρτήσεων. Η μέθοδος NEVPT2 ταυτίζεται με την μέθοδο MP2 στην περίπτωση μονο-οριζουσιακής κυματοσυνάρτησης (HF).



3.7. Ανάπτυξη CASSCF μεθοδολογίας με το διμερές $C_2H_4-H_2O$

Η πιο απλή περίπτωση ενός ασθενή δεσμού υδρογόνου με ένα π μοριακό σύστημα είναι το διμερές $C_2H_4-H_2O$ ¹⁶⁶. Το σύστημα αυτό μπορεί να χρησιμοποιηθεί ως μοντέλο για την μελέτη μεγαλύτερων π μοριακών συστημάτων¹⁶⁷ καθώς και για την ανάπτυξη νέων μεθοδολογιών. Δύο πειραματικές διαδικασίες που έγιναν¹⁶⁸ στο διμερές $C_2H_4-H_2O$ ανέδειξαν έναν ασθενή $OH..\pi$ δεσμό υδρογόνου ο οποίος διευθετείται κάθετα στο επίπεδο του C_2H_4 . Και στις δύο περιπτώσεις, τα άτομα υδρογόνου του H_2O κατευθύνονται προς το κέντρο των π μοριακών τροχιακών του C_2H_4 .

3.7.1. Η επιλογή του ενεργού χώρου CAS

Η πιο σημαντική διαδικασία προτού επιτευχθεί ένας υπολογισμός CASSCF αφορά την επιλογή του κατάλληλου ενεργού χώρου, ο οποίος αποτελείται από N ηλεκτρόνια και K μοριακά τροχιακά, όπως προαναφέρθηκε. Η επιλογή αυτή πρέπει ποιοτικά και ποσοτικά να είναι η καλύτερη όσον αφορά:

- Τον απαραίτητο αριθμό μοριακών τροχιακών K που αφορούν την μελέτη της διαμοριακής αλληλεπίδρασης.
- Τον ελάχιστο αριθμό μοριακών τροχιακών K , ώστε ο αριθμός των παραγομένων σχηματισμών CSFs να είναι ο μικρότερος.

Καθώς τα ενεργά μοριακά τροχιακά ορίζουν τον χώρο που θα αναπτυχθεί μία πλήρης CI συνδέονται άμεσα με το υπολογιστικό κόστος. Η επιλογή του ενεργού χώρου θα πρέπει να έχει την καλύτερη ποιοτική και ποσοτική συνεισφορά με το μικρότερο υπολογιστικό κόστος (μικρότερο ενεργό χώρο). Για να επιτευχθεί αυτό θα πρέπει:

- Να εκτελεστούν προκαταρκτικοί υπολογισμοί τύπου CI και να παρατηρηθεί η ποσοστιαία συνεισφορά άλλων σχηματισμών CSFs μέσω των συντελεστών τους.
- Να εκτελεστούν προκαταρκτικοί υπολογισμοί με μικρά σύνολα βάσης και να οπτικοποιηθούν τα μοριακά τροχιακά που έχουν μεγάλη συνεισφορά στην ολική ηλεκτρονική ενέργεια.

Ο πολύ-οριζουσιακός χαρακτήρας ενός μοριακού συστήματος αναδεικνύεται από τους συντελεστές της κυματοσυνάρτησης CI. Επίσης, οι αριθμοί κατάληψης των φυσικών τροχιακών που προκύπτουν ορίζουν τα μοριακά τροχιακά που θα πρέπει να συμπεριληφθούν στον ενεργό χώρο.



3.7.1.1. Η διαδικασία αναζήτησης των ενεργών μοριακών τροχιακών

Για την εύρεση των ενεργών μοριακών τροχιακών στο διμερές $C_2H_4-H_2O$ χρησιμοποιήθηκαν προκαταρκτικοί υπολογισμοί CI με μικρά σύνολα βάσεων. Η χρήση ενός μικρού συνόλου βάσης δίνει την δυνατότητα να εξερευνηθεί μεγάλο πλήθος οριζουσών σε μία κυματοσυνάρτηση CI κάνοντας χρήση όλων σχεδόν των μη-κατειλημμένων μοριακών τροχιακών. Προφανώς δεν μπορεί να χρησιμοποιηθεί πλήρης CI διότι είναι τεράστιο το πλήθος όλων των CSFs ακόμα και στην περίπτωση με το μικρότερο σύνολο βάσης STO-3G. Η διαδικασία που εφαρμόστηκε με το πακέτο GAMESS-US αναλύεται:

1. Ως αρχική γεωμετρία του $C_2H_4-H_2O$ ορίστηκε η βελτιστοποιημένη MP2/6-31+G*.
2. Επιλέχθηκε το σύνολο βάσης STO-3G. Το σύνολο STO-3G αποτελείται από 21 καρτεσιανές συναρτήσεις βάσης (μοριακά τροχιακά) για το $C_2H_4-H_2O$ (14 του C_2H_4 και 7 του H_2O). Όταν η μέθοδος RHF συγκλίνει 13 μοριακά τροχιακά θα είναι διπλά κατειλημμένα (occupied orbitals) και $21-13=8$ στοιβάδες θα αντιστοιχούν στα μη-κατειλημμένα μοριακά τροχιακά (virtual orbitals).
3. Για τον υπολογισμό τύπου CI που πρόκειται να εκτελεστεί χρειάζεται να δημιουργηθεί η κυματοσυνάρτηση CI που θα κάνει χρήση όλων των 21 μοριακών τροχιακών. Στην πράξη η κυματοσυνάρτηση αναπτύσσεται χρησιμοποιώντας τις διεγέρσεις όλων των εξωτερικών ηλεκτρονίων τα οποία είναι 20. Τα 6 εσωτερικά ηλεκτρόνια του συστήματος θεωρείται ότι πάντα καταλαμβάνουν τα 3 χαμηλότερα μοριακά τροχιακά, οπότε στην πράξη η κυματοσυνάρτηση περιέχει «κατάλληλες» διεγέρσεις των 20 ηλεκτρονίων σε 18 μοριακά τροχιακά.
4. Δημιουργείται η κυματοσυνάρτηση CI.
5. Εκτελείται ο υπολογισμός.

Όλοι οι σχηματισμοί που δημιουργούνται από 20 ηλεκτρόνια σε 18 μοριακά είναι πάρα πολλοί σύμφωνα με την εξίσωση (3.5). Για τον λόγο αυτό αναπτύσσεται μια μεθοδολογία επιλογής σχηματισμών που θα περιλαμβάνονται στην κυματοσυνάρτηση CI και περιγράφεται στην επόμενη ενότητα. Τα αποτελέσματα του υπολογισμού αναλύονται στην συνέχεια εφόσον αναπτυχθεί η διαδικασία της δημιουργίας της κυματοσυνάρτησης CI.



3.7.1.2. Δημιουργία των σχηματισμών της ανάπτυξης CI

Η κυματοσυνάρτηση CI στην οποία θα εκτελεστεί ο προκαταρκτικός υπολογισμός με το GAMESS-US πρέπει να αποτελείται από την βασική ορίζουσα Slater Φ_0 και όλες τις ορίζουσες Φ_D που αποτελούνται από διεγέρσεις τύπου $\chi_j^2 \rightarrow \chi_r^2$, όπου ένα διπλά κατειλημμένο μοριακό τροχιακό j αντικαθίσταται από ένα μη-κατειλημμένο μοριακό τροχιακό r . Δηλαδή η κυματοσυνάρτηση CI είναι της μορφής:

$$\Psi_{CI} = a_0 \Phi_0 + \sum_{D=1}^n a_D \Phi_D$$

όπου οι n όροι εξαρτώνται από τον αριθμό των εξωτερικών ηλεκτρονίων και τον αριθμό των μη-κατειλημμένων μοριακών τροχιακών.

Έτσι στην γενική περίπτωση που έχουμε N ηλεκτρόνια σε M μοριακά τροχιακά ο τύπος για το n είναι:

$$n = \frac{N}{2} \cdot \left(M - \frac{N}{2} \right)$$

και ο συνολικός αριθμός των όρων της κυματοσυνάρτησης CI είναι $n + 1$.

Στον υπολογισμό STO-3G του διμερούς $C_2H_4-H_2O$ έχουμε $N=20$ ηλεκτρόνια σε $M=18$ μοριακά τροχιακά που σημαίνει ότι η κυματοσυνάρτηση CI αποτελείται από:

$$n + 1 = \frac{20}{2} \cdot \left(18 - \frac{20}{2} \right) + 1 = 80 + 1 = 81 \text{ σχηματισμούς.}$$

Στο πακέτο GAMESS-US οι σχηματισμοί αυτοί εισάγονται ως μέρος του αρχείου εισόδου χρησιμοποιώντας το keyword \$GCILST. Οι τέσσερις πρώτες γραμμές της εισαγωγής των σχηματισμών στο αρχείο εισόδου για τον υπολογισμό του $C_2H_4-H_2O$ σε βάση STO-3G έχουν ως εξής:

```
$GCILST
81
2 2 2 2 2 2 2 2 2 2 0 0 0 0 0 0 0 0
2 2 2 2 2 2 2 2 2 0 2 0 0 0 0 0 0 0
```

Η πρώτη γραμμή είναι το keyword της λίστας και η δεύτερη ο αριθμός των σχηματισμών που θα διαβαστούν ενώ κάθε επόμενη γραμμή περιέχει έναν σχηματισμό ο οποίος αποτελείται από τα 18 μοριακά τροχιακά δεδομένου ότι τα 3 εσωτερικά μοριακά από τα 21 είναι πάντα κατειλημμένα. Ο πρώτος σχηματισμός



είναι η βασική ορίζουσα Φ_0 η οποία είναι και η βασική ορίζουσα ή ορίζουσα αναφοράς με κατειλημμένα τα $3+10=13$ χαμηλότερα μοριακά τροχιακά.

Κάθε γραμμή σχηματισμών μπορεί να πάρει τις τιμές 0, 1 ή 2 αν το τροχιακό καταλαμβάνεται από κανένα, ένα ή δύο ηλεκτρόνια αντίστοιχα. Η ενεργοποίηση του keyword \$GCILST γίνεται με την εισαγωγή της επόμενης γραμμής στο αρχείο εισόδου.

```
$CIGEN GLIST=INPUT NCORE=3 NELS=20 NACT=18 $END
```

Με αυτόν τον τρόπο δημιουργείται η κυματοσυνάρτηση CI που περιέχει τους 81 σχηματισμούς και εκτελείται ο υπολογισμός. Μιας πλήρης CI με το ίδιο σύνολο βάσης και 20 ηλεκτρόνια σε 18 μοριακά τροχιακά αποτελείται σύμφωνα με την εξίσωση (3.5) από 367,479,684 σχηματισμούς.

Για την δημιουργία των σχηματισμών που πρέπει να συμπεριληφθούν στο αρχείο εισόδου του GAMESS-US αναπτύχθηκε ένα βοηθητικό πρόγραμμα σε γλώσσα προγραμματισμού C (Παράρτημα Z) που παράγει αυτόματα όλους τους απαραίτητους σχηματισμούς. Χωρίς την χρήση της βοηθητικής αυτής εφαρμογής η είσοδος των σχηματιστών από τον χρήστη θα απαιτούσε αρκετές ώρες και τον κίνδυνο πιθανού λάθους.

3.7.1.3. Τα αποτελέσματα του υπολογισμού CI

Έχοντας σχεδιάσει τον προκαταρκτικό υπολογισμό, αναμένεται να εξαχθεί η κατάλληλη πληροφορία για τον σχεδιασμό του ενεργού χώρου της κυματοσυνάρτησης CASSCF.

Από τον υπολογισμό με την μέθοδο CI εξάγονται δύο πληροφορίες: α) το ποσοστό συμμετοχής όλων των σχηματισμών της κυματοσυνάρτησης και β) οι αριθμοί κατάληψης των φυσικών μοριακών τροχιακών.

Εκτελώντας τον παραπάνω υπολογισμό εξάγονται τα εξής αποτελέσματα για τους συντελεστές της κυματοσυνάρτησης CI που φαίνονται στο επόμενο πίνακα.

A/A	CSF	Συντελεστής
1	22222222200000000	$a_0 = 0.9662132$
2	22222222202000000	$a_1 = -0.2279100$



Η υπ' αριθμόν 1 CSF είναι η βασική ορίζουσα αναφοράς με κατειλημμένα τα χαμηλότερα 13 μοριακά τροχιακά και ποσοστό συμμετοχής $a_0^2 \approx 93\%$ (αναμένεται να έχει την μεγαλύτερη συνεισφορά), ενώ η δεύτερη CSF συμμετέχει στην κυματοσυνάρτηση με ποσοστό $a_1^2 \approx 5.1\%$. Η δεύτερη ορίζουσα με την σημαντική συνεισφορά του 5.1% στην ολική ηλεκτρονική ενέργεια προκύπτει από την προώθηση $\chi_{13}^2 \rightarrow \chi_{14}^2$, δηλαδή 2 ηλεκτρονίων από το μοριακό τροχιακό 13 στο μοριακό τροχιακό 14. Άρα υπάρχει μια ένδειξη ότι το μοριακό σύστημα $C_2H_4-H_2O$ μπορεί να αντιμετωπιστεί με πολύ-οριζουσιακού τύπου κυματοσυναρτήσεις.

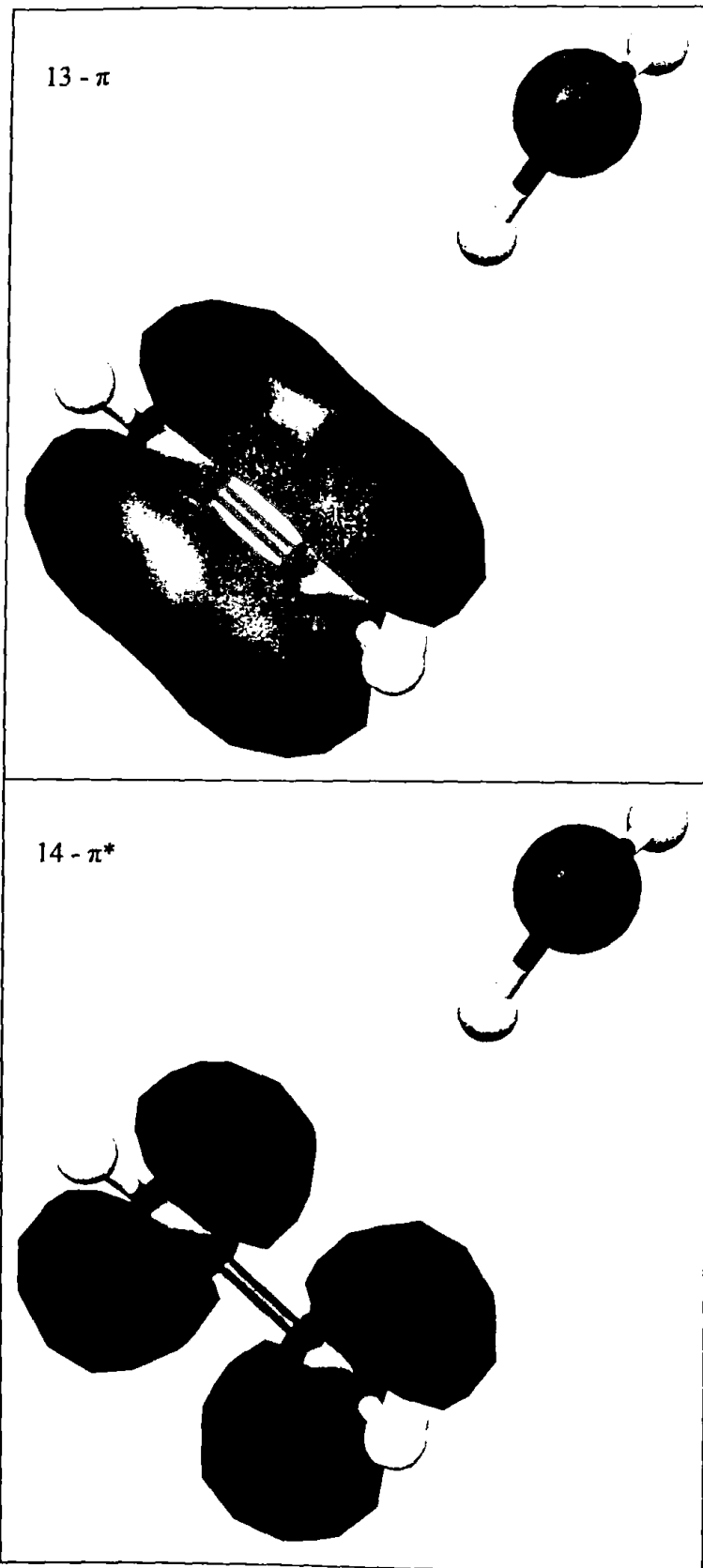
Στην συνέχεια πρέπει να εντοπιστούν τα μοριακά τροχιακά που θα αποτελέσουν τον ενεργό χώρο στον υπολογισμό CASSCF. Για να γίνει αυτό παρατηρούμε τους αριθμούς καταληψιμότητας των φυσικών μοριακών που προέκυψαν από την κυματοσυνάρτηση CI. Οι αριθμοί αυτοί κυμαίνονται από 0 για μη-κατειλημμένα μοριακά τροχιακά έως 2 για διπλά κατειλημμένα μοριακά τροχιακά. Στον επόμενο πίνακα συνοψίζονται τα αποτελέσματα καταληψιμότητας από το αρχείο εξόδου του υπολογισμού στο πακέτο GAMESS-US.

Μοριακά Τροχιακά	Πλήθος Μοριακών Τροχιακών	Αριθμοί Καταληψιμότητας
1-12	12	2.000 - 1.994
13	1	1.895
14	1	0.105
15-21	7	0.007 - 0.001

Από το πίνακα φαίνεται ότι τα 12 χαμηλότερα μοριακά τροχιακά είναι σχεδόν πάντα κατειλημμένα. Τα υψηλότερα 7 μοριακά δεν είναι σχεδόν ποτέ κατειλημμένα ενώ τα μοριακά τροχιακά 13 και 14 είναι συνήθως μερικώς κατειλημμένα. Η μερική καταληψιμότητα των τροχιακών 13 και 14 ορίζει και τον πολύ-οριζουσιακό χαρακτήρα του συστήματος όπως φάνηκε και από τους συντελεστές των σχηματισμών. Άρα ο ενεργός χώρος θα πρέπει να περιέχει τουλάχιστον τα μοριακά τροχιακά 13 και 14 τα οποία φαίνονται στην εικόνα 3.2. Αυτά τα 2 μοριακά τροχιακά είναι τα HOMO και LUMO στην συγκεκριμένη περίπτωση.



Εικόνα 3.2



Τα μοριακά τροχιακά 13 και 14 είναι τα π και π^* και όπως αναμένεται είναι αυτά στα οποία οφείλεται η ασθενής αλληλεπίδραση τύπου OH- π . Οπότε, μεγάλο μέρος της διαμοριακής ενέργειας αλληλεπίδρασης εξαρτάται από αυτά τα μοριακά τροχιακά τα οποία θα συμπεριληφθούν στον ενεργό χώρο της κυματοσυνάρτησης CASSCF για να δώσουν μια ευέλικτη πολύ-οριζουσιακή κυματοσυνάρτηση.

3.7.1.4. Επεκτείνοντας την βάση στην μέθοδο CI

Για τον υπολογισμό με κυματοσυνάρτηση CI χρησιμοποιήθηκε το μικρότερο σύνολο βάσης STO-3G. Αν χρησιμοποιηθεί ένα μεγαλύτερο σύνολο βάσης τα συνολικά μοριακά τροχιακά θα αυξηθούν, οπότε η προηγούμενη διαδικασία θα πρέπει να εφαρμοστεί όπως περιγράφεται στην συνέχεια.

Επιλέγοντας το σύνολο βάσης 3-21G, το οποίο είναι γενικά ένα μικρό σύνολο βάσης, το διμερές $C_2H_4-H_2O$ αποτελείται από 39 συναρτήσεις βάσης κλειστών στοιβάδων (μοριακά τροχιακά). Εφαρμόζοντας την προηγούμενη διαδικασία με 20 ηλεκτρόνια σε 30 μοριακά τροχιακά αναπτύσσεται μία κυματοσυνάρτηση CI με 201 όρους. Τα μοριακά τροχιακά 13 και 14 είναι μερικώς κατελημμένα ενώ η βασική ορίζουσα έχει ποσοστό συμμετοχής 95.7% και ο επόμενος σχηματισμός έχει ποσοστό συμμετοχής 3.3%. Κατά αναλογία ο σχηματισμός που συνεισφέρει στην ολική ηλεκτρονική ενέργεια εκτός της βασικής ορίζουσας προέρχεται από την $\chi_{13}^2 \rightarrow \chi_{14}^2$ προώθηση.

Η χρήση του συνόλου βάσης 6-31+G*, το οποίο έχει 69 μοριακά τροχιακά για το διμερές $C_2H_4-H_2O$, καταλήγει στα ίδια ενεργά μοριακά τροχιακά.

3.7.1.5. Ο ενεργός χώρος για το διμερές $C_2H_4-H_2O$

Από τους προκαταρκτικούς υπολογισμούς με χρήση της κυματοσυνάρτησης CI φαίνεται ότι ο ενεργός χώρος πρέπει να αποτελείται από τα μοριακά τροχιακά π και π^* της εικόνας 3.2. Επειδή στους μοριακούς κβαντικούς υπολογισμούς τα μοριακά τροχιακά ταξινομούνται κατά την ενέργεια τους και καθότι η χρήση μεγάλων συνόλων βάσεων είναι απαραίτητη για υπολογισμούς με μεγάλη ακρίβεια τα μοριακά τροχιακά π και π^* ενδέχεται να έχουν διαφορετική αρίθμηση με διαφορετικά σύνολα βάσεων. Σε όλους τους υπολογισμούς CASSCF που έγιναν το μοριακό τροχιακό π είναι πάντα το HOMO ενώ δεν ισχύει το ίδιο για το π^* , το οποίο είναι LUMO+n, με το n να εξαρτάται από το σύνολο βάσης και να χρειάζεται οπτικοποίηση για τον εντοπισμό του.



Για τον υπολογισμό της διαμοριακής ενέργειας αλληλεπίδρασης με την μέθοδο CASSCF στο διμερές $C_2H_4-H_2O$, ο ενεργός χώρος πρέπει να αποτελείται από τα δύο μοριακά τροχιακά π και π^* της εικόνας 3.2. Άρα ο υπολογισμός θα γίνει με $n=2$ ηλεκτρόνια και $m=2$ μοριακά τροχιακά που σημαίνει CASSCF(2,2).

Συνοψίζοντας, βλέπουμε ότι με την εφαρμογή μιας κατάλληλης μεθοδολογίας βρίσκουμε τον κατάλληλο ενεργό χώρο που θα χρησιμοποιηθεί για τους υπολογισμούς με την μέθοδο CASSCF.

3.7.1.6. Επεκτείνοντας τον ενεργό χώρο

Η επέκταση του ενεργού χώρου γενικά δίνει ακρίβεια και ευελιξία στους υπολογισμούς CASSCF. Όμως ο ενεργός χώρος πρέπει να αποτελείται από κατάλληλα μοριακά τροχιακά ώστε να συνεισφέρουν στην ολική ηλεκτρονική ενέργεια και στην μελέτη του μοριακού συστήματος. Ο υπολογισμός CASSCF(2,2) αποτελείται από ενεργά μοριακά τροχιακά μόνο του C_2H_4 . Μπορούμε να δώσουμε ευελιξία και να αυξήσουμε τον ενεργό χώρο χρησιμοποιώντας και ενεργά μοριακά τροχιακά από το H_2O . Αυτό θα οδηγήσει σε μεγαλύτερο ενεργό χώρο.

Παρατηρώντας τους αριθμούς καταληψιμότητας των φυσικών μοριακών από τον υπολογισμό τύπου CI/STO-3G βλέπουμε ότι το μοριακό τροχιακό 15 μπορεί να προστεθεί στον ενεργό. Το ίδιο επιβεβαιώνεται και από τους άλλους υπολογισμούς. Η καταληψιμότητα του μοριακού τροχιακού είναι 0.0079 οπότε μπορεί να συμπεριληφθεί στον ενεργό χώρο.

Γνωρίζουμε όμως ότι ο καλύτερος υπολογισμός CASSCF είναι αυτός που συμπεριλαμβάνει στον ενεργό χώρο όλα τα εξωτερικά ηλεκτρόνια, το οποίο είναι πρακτικά αδύνατο να εφαρμοστεί στις περισσότερες περιπτώσεις λόγω του τεράστιου αριθμού σχηματισμών που θα προέκυπταν από την εξίσωση (3.5).

Για την μελέτη της διαμοριακής ενέργειας αλληλεπίδρασης είναι δυνατόν να χρησιμοποιηθεί ενεργός χώρος με όλα τα εξωτερικά ηλεκτρόνια του H_2O συμπεριλαμβανομένου του LUMO μοριακού τροχιακού του, το οποίο συμπίπτει με το προαναφερθέν μοριακό τροχιακό 15. Ο ενεργός χώρος για τον υπολογισμό του H_2O είναι $n=8$ και $m=5$ και η κυματοσυνάρτηση CASSCF(8,5).

Προσθέτοντας τον ενεργό χώρο CASSCF(8,5) του H_2O με τον ενεργό χώρο CASSCF(2,2) που προέκυψε αρχικά, έχουμε έναν νέο ενεργό χώρο CASSCF(10,7) για την μελέτη της διαμοριακής ενέργειας αλληλεπίδρασης στο διμερές $C_2H_4-H_2O$.



3.7.1.7. Size-extensivity στον ενεργό χώρο

Η επιλογή των μεθόδων CASSCF(2,2) και CASSCF(10,7) για τον υπολογισμό των ολικών ηλεκτρονικών ενεργειών καθώς και των ενεργειών διαμοριακής αλληλεπίδρασης του $C_2H_4-H_2O$ θα πρέπει να ακολουθεί τους κανόνες size-extensivity ή size-consistency όπως έχει συνηθιστεί σε υπολογισμούς διμερών. Στην περίπτωση του υπολογισμού CASSCF(2,2) και σε μεγάλη διαμοριακή απόσταση R θα πρέπει:

$$E_{C_2H_4-H_2O}^R [CASSCF(2,2)] = E_{C_2H_4} [CASSCF(2,2)] + E_{H_2O} [RHF]$$

ενώ για τους υπολογισμούς CASSCF(10,7) θα πρέπει:

$$E_{C_2H_4-H_2O}^R [CASSCF(10,7)] = E_{C_2H_4} [CASSCF(2,2)] + E_{H_2O} [CASSCF(8,5)]$$

Ο εκθέτης R συμβολίζει την σχετική διαμοριακή απόσταση από την θέση ισορροπίας ($R=0$). Το $R=0$ αντιστοιχεί στην βελτιστοποιημένη γεωμετρική διεύθετη σε δεδομένο σύνολο βάσης.

Η ισότητα σε κάθε μία από τις παραπάνω σχέσεις θα επαληθευθεί στην συνέχεια. Η χρήση αυτών των σχέσεων δίνει την δυνατότητα για την εφαρμογή της μεθόδου CP (counterpoise, κεφάλαιο 2) των Boys & Bernardi, για τον υπολογισμό της διαμοριακής ενέργειας αλληλεπίδρασης διορθωμένης ως προς το BSSE, σε κυματοσυναρτήσεις CASSCF.

3.7.1.8. Η υπολογιστική μεθοδολογία

Τα βήματα για την υπολογιστική μεθοδολογία που αναπτύχθηκε για την μελέτη της διαμοριακής ενέργειας αλληλεπίδρασης ετεροδιμερών έχει ως εξής:

1. Γίνεται βελτιστοποίηση του ετεροδιμερούς χρησιμοποιώντας την θεωρία MP2 στο σύνολο βάσης B.
2. Στην MP2/B βελτιστοποιημένη γεωμετρία, γίνεται προκαταρκτικός CI/STO-3G (με διπλές διεγέρσεις) υπολογισμός για να εντοπιστούν τα π τροχιακά που θα συμπεριληφθούν στον ενεργό χώρο. Οποιαδήποτε άλλη βιβλιογραφική μέθοδος εντοπισμού τροχιακών μπορεί να χρησιμοποιηθεί.
3. Γίνεται RHF υπολογισμός με τον σύνολο βάσης B στην MP2/B βελτιστοποιημένη γεωμετρία.
4. Γίνεται οπτικοποίηση των βελτιστοποιημένων μοριακών τροχιακών από τα βήματα 2 και 3 για να κατασκευαστεί ο εσωτερικός (core) και ενεργός (active) χώρος.



5. Γίνεται CASSCF/B υπολογισμός, χρησιμοποιώντας ως αρχικά, τα βελτιστοποιημένα μοριακά τροχιακά στο επίπεδο RHF/B. Σε αυτό το σημείο τα μοριακά τροχιακά χρειάζεται να αναδιαταχθούν στις περισσότερες περιπτώσεις.
6. Σε μεγάλη διαμοριακή απόσταση της MP2/B βελτιστοποιημένης γεωμετρίας γίνεται CASSCF/B υπολογισμός, χρησιμοποιώντας ως αρχικά τα βελτιστοποιημένα RHF/B μοριακά τροχιακά.
7. Επαναλαμβάνονται τα βήματα 3,4,5 δύο φορές για κάθε μονομερές στον δικό του ενεργό χώρο.
8. Ελέγχεται η size-consistency με υπολογισμό CASSCF/B χρησιμοποιώντας τα αποτελέσματα από τα βήματα 6 and 7.
9. Εφόσον η size-consistency από το βήμα 8 αποτύχει, τότε τα εσωτερικά (core) και τα ενεργά (active) τροχιακά των μονομερών δεν ταυτίζονται με αυτά που χρησιμοποιήθηκαν στο βήμα 5 και η διαδικασία πρέπει να επαναληφθεί από το βήμα 4.
10. Εφόσον η size-consistency από το βήμα 8 επαληθευθεί, μπορούμε να συνεχίσουμε τους υπολογισμούς (π.χ. NEVPT2) και να λάβουμε τα αποτελέσματα.

Μια τεχνική οδηγία που μπορεί να ληφθεί υπ' όψη, αφορά το βήμα 5 της διαδικασίας, όπου τα βελτιστοποιημένα CASSCF/B μοριακά τροχιακά μπορούν να χρησιμοποιηθούν σε επόμενους υπολογισμούς χωρίς αναδιάταξη. Για τις εφαρμογές που θα δούμε στην συνέχεια αυτό έγινε επιτυχώς και μεταφέρθηκαν τα βελτιστοποιημένα μοριακά τροχιακά CASSCF/B από το GAMESS-US στο πακέτο Dalton για τους υπολογισμούς με την μέθοδο NEVPT2.



3.7.2. Ο ενεργός χώρος CASSCF(2,2)

Ο ενεργός χώρος CASSCF(2,2) του $C_2H_4-H_2O$ αποτελείται από $N=2$ ηλεκτρόνια και $K=2$ μοριακά τροχιακά. Από την εξίσωση (3.5) μπορούμε να υπολογίσουμε τον αριθμό των συνολικών σχηματισμών που συμμετέχουν στην κυματοσυνάρτηση CASSCF(2,2) με ολικό spin $S=0$.

$$CSFs = \frac{2 \cdot 0 + 1}{2 + 1} \cdot \binom{2 + 1}{2 - 1/2 \cdot 2 - 0} \cdot \binom{2 + 1}{1/2 \cdot 2 - 0} = \frac{1}{3} \cdot \binom{3}{1} \cdot \binom{3}{1} = \frac{1}{3} \cdot 3 \cdot 3 = 3$$

Άρα η πολύ-οριζουσιακή κυματοσυνάρτηση CASSCF(2,2) είναι της μορφής:

$$\Psi_{CASSCF} = \sum_{i=1}^3 a_i \Phi_i, \text{ όπου}$$

$$\Phi_i = \sum_{j=1}^2 k_j \chi_j, \text{ με την προϋπόθεση } \sum_{i=1}^3 a_i^2 = 1$$

και ζητούνται τα βέλτιστα a_i και k_j που ελαχιστοποιούν την ολική ηλεκτρονική ενέργεια υπό την επίδραση του χαμιλτωνιανού τελεστή στην κυματοσυνάρτηση.

3.7.2.1. Υπολογισμός της ενέργειας CASSCF(2,2)/6-31+G*

Ο πρώτος υπολογισμός της ολικής ηλεκτρονικής ενέργειας του διμερούς $C_2H_4-H_2O$ με την κυματοσυνάρτηση CASSCF(2,2) και το σύνολο βάσης 6-31+G* γίνεται στην βελτιστοποιημένη MP2/6-31+G* γεωμετρία.

Αρχικά, στην γεωμετρία αυτή εκτελείται ένας ενεργειακός υπολογισμός RHF/6-31+G* και προκύπτουν τα 69 βελτιστοποιημένα RHF μοριακά τροχιακά. Τα πρώτα 13 διπλά κατειλημμένα μοριακά τροχιακά είναι αυτά που ορίζουν την ενέργεια της βασικής κατάστασης του συστήματος $C_2H_4-H_2O$.

Για τον ενεργό χώρο της CASSCF(2,2) χρειάζεται δύο μοριακά τροχιακά να συμπεριληφθούν στον ενεργό χώρο. Αυτά είναι τα π και π^* της εικόνας 3.2. Ο υπολογισμός CASSCF(2,2) θα εκτελεστεί θεωρώντας ότι 12 μοριακά τροχιακά είναι πάντα διπλά κατειλημμένα (core orbitals) και 2 μοριακά τροχιακά είναι ενεργά. Αυτά τα 2 ενεργά μοριακά τροχιακά εντοπίζονται με την διαδικασία της οπτικοποίησης από το σύνολο των 69 και είναι: α) το π στη θέση 13 και β) το π^* στην θέση 16.

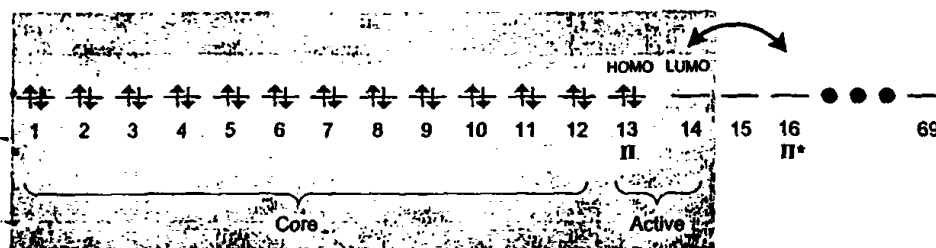
Ο υπολογισμός CASSCF(2,2)/6-31+G* στο πακέτο GAMESS-US θα πρέπει να ξεκινήσει με αρχικά μοριακά τροχιακά τα βελτιστοποιημένα RHF/6-31+G*.

Το μοριακό τροχιακό 16 που θα συμπεριληφθεί στον ενεργό χώρο πρέπει να αλλάξει θέση και να πάει στην θέση 14 ώστε ο ενεργός χώρος να αποτελείται από τα



επιθυμητά μοριακά τροχιακά 13 και 14, δηλαδή το HOMO και LUMO. Αυτό φαίνεται στην εικόνα 3.3.

Εικόνα 3.3



Γενικά σε υπολογισμούς CASSCF με το πακέτο GAMESS-US ο ενεργός χώρος πρέπει να είναι «συνεχόμενος» από HOMO-m έως LUMO+n και όλα τα διπλά κατειλημμένα μοριακά τροχιακά να είναι από 1 έως HOMO-m-1 (εικόνα 3.3). Στον ενεργό χώρο CASSCF(2,2) το m=0 και n=0.

Ο ορισμός του ενεργού χώρου στο πακέτο GAMESS-US φαίνεται στην επόμενη γραμμή:

```
$DET NCORE=12 NACT=2 NELS=2 $END
```

Η είσοδος των μοριακών τροχιακών από τον υπολογισμό RHF/6-31+G* και η αναδιάταξη των μοριακών τροχιακών 14 και 16 γίνεται στο GAMESS-US όπως φαίνεται στην επόμενη γραμμή:

```
$guess guess=MOREAD NORB=69 NORDER=1 IORDER(14)=16,15,14 $end
```

Ο υπολογισμός CASSCF(2,2)/6-31+G* εκτελείται και μετά την ολοκλήρωση της επαναληπτικής διαδικασίας η ολική ηλεκτρονική ενέργεια είναι:

$$E_{C_{2H_4-H_2O}}^0 [CASSCF(2,2)] = -154.0812446612 \text{ a.u.}$$

ενώ η RHF/6-31+G* ενέργεια είναι:

$$E_{C_{2H_4-H_2O}}^0 [RHF] = -154.0535845335 \text{ a.u.}$$

Η ενεργειακή διαφορά των δύο μεθόδων είναι:

$$E_{C_{2H_4-H_2O}}^0 [CASSCF(2,2)] - E_{C_{2H_4-H_2O}}^0 [RHF] = -17.6 \text{ Kcal/mol}$$

ποσότητα που ως επί το πλείστον αντιστοιχεί στην μη-δυναμική ενέργεια ηλεκτρονικής συσχέτισης του συστήματος C₂H₄-H₂O στο σύνολο βάσης 6-31+G*.

3.7.2.2. Ο υπολογισμός της CASSCF(2,2)/6-31+G* διαμοριακής ενέργειας αλληλεπίδρασης



Ο υπολογισμός της διαμοριακής ενέργειας αλληλεπίδρασης θα πρέπει να επιβεβαιώνει το κριτήριο της size-consistency στον υπολογισμό CASSCF(2,2)/6-31+G*, δηλαδή την εξίσωση:

$$E_{C_2H_4-H_2O}^R [CASSCF(2,2)] = E_{C_2H_4} [CASSCF(2,2)] + E_{H_2O} [RHF]$$

για μεγάλη διαμοριακή απόσταση R (κεφ 2).

Ο υπολογισμός γίνεται εφαρμόζοντας την διαδικασία της ενότητας 3.7.2.1, θέτοντας την διαμοριακή απόσταση $R=200\text{\AA}$, ώστε να μην υπάρχει διαμοριακή αλληλεπίδραση. Η γεωμετρία του $C_2H_4-H_2O$ στο αρχείο εισόδου του GAMESS-US είναι η επόμενη:

C 6.0	0.6648609717	-0.0059289223	-0.0027188005
C 6.0	-0.6757460905	0.0025791796	-0.0018139283
H 1.0	1.2383773806	0.7601245941	0.5103957269
H 1.0	1.2277913110	-0.7847428229	-0.5081897484
H 1.0	-1.2492240151	-0.7689114682	-0.5066420571
H 1.0	-1.2388386832	0.7759349339	0.5119042340
O 8.0	0.4756140089	175.8486974475	-102.1096585861
H 1.0	0.4993683734	175.8337322640	-103.0798036518
H 1.0	0.4750565723	174.9159406039	-101.8314081070

Σε αυτή την γεωμετρία εκτελείται ο υπολογισμός CASSCF(2,2)/6-31+G* με αρχικά μοριακά τροχιακά αυτά που χρησιμοποιηθήκαν για τον υπολογισμό της ενέργειας $E_{C_2H_4-H_2O}^0 [CASSCF(2,2)]$.

Η ολική ηλεκτρονική ενέργεια του $C_2H_4-H_2O$ από τον υπολογισμό CASSCF(2,2)/6-31+G* είναι:

$$E_{C_2H_4-H_2O}^{200} [CASSCF(2,2)] = -154.0796675 \text{ a.u.}$$

Ο επόμενος υπολογισμός πρέπει να γίνει για το H_2O του διμερούς συστήματος. Ο υπολογισμός αυτός είναι RHF/6-31+G* και η ολική ηλεκτρονική ενέργεια είναι:

$$E_{H_2O} [RHF] = -76.01653735 \text{ a.u.}$$

Ο τελικός επιμέρους υπολογισμός είναι ο CASSCF(2,2)/6-31+G* για το C_2H_4 . Ο ενεργός χώρος αποτελείται από τα δύο π και π^* μοριακά τροχιακά (εικόνας 3.2) του C_2H_4 . Η διαδικασία που ακολουθείτε περιγράφεται στην ενότητα 3.7.2.1. Η ολική ηλεκτρονική ενέργεια που προκύπτει είναι:



$$E_{\text{C}_2\text{H}_4} [\text{CASSCF}(2,2)] = -78.06313015 \text{ a.u.}$$

Προσθέτοντας τις επιμέρους ολικές ηλεκτρονικές ενέργειες των C_2H_4 και H_2O έχουμε:

$$E_{\text{C}_2\text{H}_4} [\text{CASSCF}(2,2)] + E_{\text{H}_2\text{O}} [\text{RHF}] = -78.06313015 - 76.01653735 = -154.0796675 \text{ a.u.}$$

η οποία είναι ακριβώς ίδια με την:

$$E_{\text{C}_2\text{H}_4\text{-H}_2\text{O}}^{200} [\text{CASSCF}(2,2)] = -154.0796675 \text{ a.u.}$$

Άρα επιβεβαιώνεται ακριβώς η size-consistency του συστήματος στους υπολογισμούς CASSCF(2,2)/6-31+G*. Η διαμοριακή ενέργεια αλληλεπίδρασης του $\text{C}_2\text{H}_4\text{-H}_2\text{O}$ με το σύνολο βάσης 6-31+G* και την κυματοσυνάρτηση CASSCF(2,2) χρησιμοποιώντας την μέθοδο υπερμοριακής προσέγγισης ορίζεται ως:

$$\delta E_{\text{C}_2\text{H}_4\text{-H}_2\text{O}} = E_{\text{C}_2\text{H}_4\text{-H}_2\text{O}}^0 [\text{CASSCF}(2,2)] - E_{\text{C}_2\text{H}_4\text{-H}_2\text{O}}^{200} [\text{CASSCF}(2,2)]$$

και είναι:

$$\Delta E_{\text{C}_2\text{H}_4\text{-H}_2\text{O}} = -0.99 \text{ Kcal / mol}$$

Ο έλεγχος και η επιβεβαίωση της size-consistency αποτελεί και ένα κριτήριο σωστής σύγκλισης της CASSCF σε διμερή μοριακά συστήματα.

Στις επόμενες ενότητες θα εφαρμοστεί η διαδικασία counterpoise για τον υπολογισμό της διαμοριακής ενέργειας αλληλεπίδρασης χρησιμοποιώντας την μέθοδο CASSCF και θα γίνει η αξιολόγηση των αποτελεσμάτων.



3.7.3. Ο ενεργός χώρος CASSCF(10,7)

Ο ενεργός χώρος CASSCF(10,7) του $C_2H_4-H_2O$ αποτελείται από $N=10$ ηλεκτρόνια και $K=7$ μοριακά τροχιακά. Από την εξίσωση (3.5) μπορούμε να υπολογίσουμε τον αριθμό των συνολικών σχηματισμών που συμμετέχουν στην κυματοσυνάρτηση CASSCF(10,7) με ολικό spin $S=0$.

$$CSFs = \frac{2 \cdot 0 + 1}{7 + 1} \cdot \binom{7 + 1}{7 - 1/2 \cdot 10 - 0} \cdot \binom{7 + 1}{1/2 \cdot 10 - 0} = \frac{1}{8} \binom{8}{2} \cdot \binom{8}{5} = \frac{1}{8} \cdot 28 \cdot 56 = 196$$

Άρα η πολύ-οριζουσιακή κυματοσυνάρτηση CASSCF(10,7) είναι της μορφής:

$$\Psi_{CASSCF} = \sum_{i=1}^{196} a_i \Phi_i, \text{ όπου}$$

$$\Phi_i = \sum_{j=1}^7 k_j \chi_j, \text{ με την προϋπόθεση } \sum_{i=1}^{196} a_i^2 = 1$$

και ζητούνται τα βέλτιστα a_i και k_j που ελαχιστοποιούν την ολική ηλεκτρονική ενέργεια υπό την επίδραση του χαμιλτωνιανού τελεστή στην κυματοσυνάρτηση.

Ο ενεργός χώρος CASSCF(10,7) είναι αρκετά μεγαλύτερος από τον CASSCF(2,2) που περιγραφητέ. Η εφαρμογή του CASSCF(10,7) στο διμερές $C_2H_4-H_2O$ πρέπει να έχει την ίδια ποσοτική συμπεριφορά με τον χώρο CASSCF(2,2). Με την CASSCF(10,7) δίνεται μεγαλύτερη ευελιξία στον ενεργό χώρο ενώ αυξάνονται αρκετά και οι σχηματισμοί (CSFs) από 3 σε 196.

3.7.3.1. Υπολογισμός της ενέργειας CASSCF(10,7)/6-31+G*

Η διαδικασία για τον υπολογισμό της ολικής ηλεκτρονικής ενέργειας του $C_2H_4-H_2O$ με την μέθοδο CASSCF(10,7) και το σύνολο βάσης 6-31+G* είναι αντίστοιχη με αυτή που περιγράφηκε στην ενότητα 3.7.2.1.

Ο ενεργός χώρος θα αποτελείται από 7 ενεργά μοριακά τροχιακά και 10 ηλεκτρόνια. Από τα 7 ενεργά μοριακά τροχιακά 2 είναι τα π και π^* του C_2H_4 και 5 είναι του H_2O . Η επιλογή των ενεργών μοριακών τροχιακών πρέπει να γίνει από τα 69 βελτιστοποιημένα μοριακά τροχιακά του υπολογισμού RHF/6-31+G*.

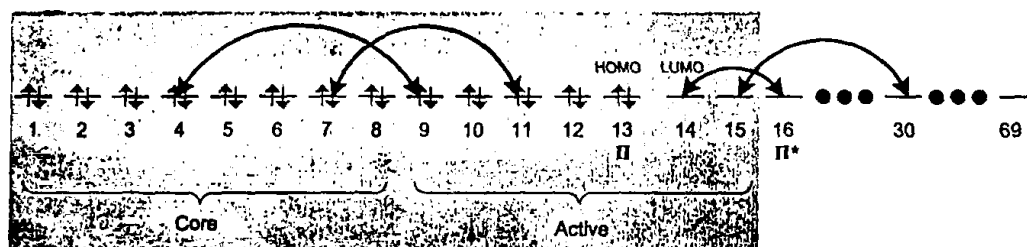
Το HOMO είναι το μοριακό τροχιακό 13 και το LUMO είναι το μοριακό τροχιακό 14. Τα ενεργά μοριακά τροχιακά θα είναι: από HOMO-4 έως LUMO+1, που σημαίνει ότι τα 8 μοριακά τροχιακά 1 έως HOMO-5 θα είναι πάντα διπλά κατειλημμένα (core). Το LUMO θα είναι το μοριακό τροχιακό π^* ενώ το LUMO+1 θα είναι το μοριακό τροχιακό του H_2O που αναφέρθηκε στην ενότητα 3.7.1.6.



Για τον υπολογισμό CASSCF(10,7) πρέπει να γίνει ο κατάλληλος εντοπισμός των μοριακών τροχιακών μέσα στα 69 βελτιστοποιημένα από τον υπολογισμό RHF/6-31+G*.

Τα 2 ενεργά μοριακά τροχιακά π και π^* προέρχονται από το C_2H_4 και είναι τα 13 και 16 όπως είδαμε προηγουμένως (εικόνα 3.2).

Εικόνα 3.4



Τα 5 ενεργά μοριακά τροχιακά που πρέπει να εντοπιστούν ανήκουν στο H_2O . Τα 4 από αυτά θα αναζητηθούν μέσα στα χαμηλότερα 13 του υπολογισμού RHF/6-31+G* για το διμερές $C_2H_4-H_2O$. Το ένα μοριακό τροχιακό που υπολείπεται θα αναζητηθεί στο σύνολο των μη-κατειλημμένων (virtual) του ίδιου υπολογισμού. Ξέρουμε την μορφή του τροχιακού αυτού από την ενότητα 3.7.1.6, οπότε θα χρησιμοποιήσουμε την διαδικασία της οπτικοποίησης για να το εντοπίσουμε. Παρατηρώντας τα μοριακά τροχιακά βλέπουμε ότι το μοριακό τροχιακό που θα συμπεριληφθεί στον ενεργό χώρο είναι το 30.

Τα 4 μοριακά τροχιακά του H_2O που πρέπει να εντοπίσουμε στα 13 χαμηλότερα από τα 69 του υπολογισμού RHF/6-31+G* του διμερούς είναι τα 4, 7, 10 και 12.

Συνοψίζοντας, ο ενεργός χώρος για τον υπολογισμό CASSCF(10,7)/6-31+G* θα αποτελείται από τα 7 μοριακά τροχιακά 4,10,7,12,13,16,30 όπως φαίνονται στην εικόνα 3.4 μετά την κατάλληλη αναδιάταξη. Ο ορισμός του ενεργού χώρου στο πακέτο GAMESS-US γίνεται ως εξής:

```
$DET NCORE=8 NACT=7 NELS=10 $END
```

Η είσοδος των μοριακών τροχιακών από τον υπολογισμό RHF/6-31+G* και η αναδιάταξη των μοριακών τροχιακών γίνεται όπως φαίνεται ακολούθως:

```
$guess guess=MOREAD NORB=69 NORDER=1
IORDER(4)=9,5,6,11,8,4,10,7,12,13,16,30,14
IORDER(30)=15
$end
```



Ο υπολογισμός CASSCF(10,7)/6-31+G* εκτελείται και μετά την ολοκλήρωση της επαναληπτικής διαδικασίας η ολική ηλεκτρονική ενέργεια είναι:

$$E_{C_2H_4-H_2O}^0 [CASSCF(10,7)] = -154.1038609180 \text{ a.u.}$$

ενώ η RHF/6-31+G* ενέργεια του διμερούς είναι:

$$E_{C_2H_4-H_2O}^0 [RHF] = -154.0535845335 \text{ a.u.}$$

Η ενεργειακή διαφορά των δύο μεθόδων είναι:

$$E_{C_2H_4-H_2O}^0 [CASSCF(10,7)] - E_{C_2H_4-H_2O}^0 [RHF] = -31.5 \text{ Kcal / mol}$$

ποσότητα που ως επί το πλείστον αντιστοιχεί στην μη-δυναμική ενέργεια ηλεκτρονικής συσχέτισης του συστήματος $C_2H_4-H_2O$ στο σύνολο βάσης 6-31+G*.

Η αύξηση του ενεργού χώρου αντιστοιχεί στην ποιοτική βελτίωση την ολικής ηλεκτρονικής ενέργειας του διμερούς, όπως αναμένεται. Σε σύγκριση με τον υπολογισμό CASSCF(2,2)/6-31+G* η ολική ηλεκτρονική ενέργεια ελαττώνεται κατά -13.9 Kcal / mol .

3.7.3.2. Ο υπολογισμός της CASSCF(10,7)/6-31+G* διαμοριακής ενέργειας αλληλεπίδρασης

Ο υπολογισμός της διαμοριακής ενέργειας αλληλεπίδρασης θα πρέπει να επιβεβαιώνει το κριτήριο της size-consistency στον υπολογισμό CASSCF(10,7)/6-31+G*, δηλαδή την εξίσωση:

$$E_{C_2H_4-H_2O}^R [CASSCF(10,7)] = E_{C_2H_4} [CASSCF(2,2)] + E_{H_2O} [CASSCF(8,5)]$$

για μεγάλη διαμοριακή απόσταση R (κεφ 2).

Ο υπολογισμός γίνεται εφαρμόζοντας την διαδικασία της ενότητας 3.7.2.1, θέτοντας την διαμοριακή απόσταση $R=200\text{\AA}$, ώστε να μην υπάρχει διαμοριακή αλληλεπίδραση. Η γεωμετρία του $C_2H_4-H_2O$ στο αρχείο εισόδου του GAMESS-US αναφέρεται στην ενότητα 3.7.2.2.

Σε αυτή την γεωμετρία εκτελείται ο υπολογισμός CASSCF(10,7)/6-31+G* με αρχικά μοριακά τροχιακά αυτά που χρησιμοποιήθηκαν για τον υπολογισμό της ενέργειας $E_{C_2H_4-H_2O}^0 [CASSCF(10,7)]$.

Η ολική ηλεκτρονική ενέργεια του $C_2H_4-H_2O$ από τον υπολογισμό CASSCF(10,7)/6-31+G* είναι:

$$E_{C_2H_4-H_2O}^{200} [CASSCF(10,7)] = -154.10221760 \text{ a.u.}$$



Ο επόμενος υπολογισμός πρέπει να γίνει για το μόριο του H_2O του διμερούς συστήματος. Ο υπολογισμός αυτός είναι CASSCF(8,5)/6-31+G* και η ολική ηλεκτρονική ενέργεια είναι:

$$E_{H_2O}[CASSCF(8,5)] = -76.0390875 \text{ a.u.}$$

Ο τελικός επιμέρους υπολογισμός είναι ο CASSCF(2,2)/6-31+G* για το C_2H_4 όπως ακριβώς περιγράφηκε στην ενότητα 4.5.2.2. Η ολική ηλεκτρονική ενέργεια είναι:

$$E_{C_2H_4}[CASSCF(2,2)] = -78.0631301 \text{ a.u.}$$

Προσθέτοντας τις επιμέρους ολικές ηλεκτρονικές ενέργειες των C_2H_4 και H_2O έχουμε:

$$E_{C_2H_4}[CASSCF(2,2)] + E_{H_2O}[CASSCF(8,5)] = -78.0631301 - 76.0390875 = -154.1022176 \text{ a.u.}$$

η οποία είναι ακριβώς ίδια με την:

$$E_{C_2H_4-H_2O}^{200}[CASSCF(10,7)] = -154.1022176 \text{ a.u.}$$

Άρα επιβεβαιώνεται ακριβώς η size-consistency του συστήματος στους υπολογισμούς CASSCF(10,7)/6-31+G*. Η διαμοριακή ενέργεια αλληλεπίδρασης του $C_2H_4-H_2O$ με το σύνολο βάσης 6-31+G* και την κυματοσυνάρτηση CASSCF(2,2) χρησιμοποιώντας την μέθοδο υπερμοριακής προσέγγισης ορίζεται ως:

$$\delta E_{C_2H_4-H_2O} = E_{C_2H_4-H_2O}^0[CASSCF(10,7)] - E_{C_2H_4-H_2O}^{200}[CASSCF(10,7)]$$

και είναι:

$$\Delta E_{C_2H_4-H_2O} = -1.03 \text{ Kcal/mol}$$

Η τιμή αυτή είναι σχεδόν ίδια με αυτή που προέκυψε από τους υπολογισμούς CASSCF(2,2)/6-31+G* και βρέθηκε -0.99 Kcal/mol . Στην συνέχεια θα εφαρμοστεί η μέθοδος των Boys-Bernardi για την απαλοιφή του BSSE από την διαμοριακή ενέργεια αλληλεπίδρασης του συστήματος $C_2H_4-H_2O$.



3.8. Απαλοιφή του BSSE σε υπολογισμούς CASSCF/6-31+G* για το C₂H₄-H₂O

Στις ενότητες 3.7.2.2 και 3.7.3.2 χρησιμοποιήθηκε η μέθοδος της υπερμοριακής προσέγγισης για τον υπολογισμό της διαμοριακής ενέργειας αλληλεπίδρασης του διμερούς C₂H₄-H₂O χρησιμοποιώντας δύο διαφορετικούς ενεργούς χώρους με την μέθοδο CASSCF. Η ενεργειακή περιγραφή του διαμοριακού συστήματος βελτιώθηκε ποιοτικά με την χρήση της μεθόδου CASSCF, αλλά οι ενεργοί χώροι που χρησιμοποιήθηκαν είναι σαφώς μικρότεροι από την πλήρη CI η οποία θα έδινε την «ακριβή» ενεργειακή περιγραφή του συστήματος. Για τον λόγο αυτό η διαμοριακή ενέργεια αλληλεπίδρασης δεν είναι απαλλαγμένη από την ποσότητα BSSE η οποία δίνεται από την επόμενη εξίσωση (βλ. κεφάλαιο 2) και είναι:

$$\delta^{BSSE} = E_A(AB) - E_A(A) + E_B(AB) - E_B(B)$$

όπου το A συμβολίζει το C₂H₄ και το B συμβολίζει το H₂O.

Τα δύο νέα προβλήματα που τίθενται και θα αντιμετωπιστούν στην συνέχεια αφορούν:

- Την εφαρμογή της μεθόδου CP (counterpoise) των Boys και Bernardi σε κυματοσυναρτήσεις CASSCF.
- Την ποσοτική μελέτη της διαμοριακής ενέργειας αλληλεπίδρασης απαλείφοντας την ποσότητα BSSE.

Η μέθοδος CP θα εφαρμοστεί για τις μεθόδους CASSCF(2,2)/6-31+G* και CASSCF(10,7)/6-31+G*. Η size-consistency της CASSCF αποδείχτηκε προηγουμένως για τους 2 ενεργούς χώρους που μελετώνται και απομένουν οι υπολογισμοί με συναρτήσεις “ghost”.

3.8.1 Υπολογισμός του BSSE με την μέθοδο CASSCF(2,2)/6-31+G*

Στην μέθοδο CASSCF(2,2)/6-31+G* θα υπολογιστεί η ενέργεια του C₂H₄ με συναρτήσεις “ghost” στο μόριο του H₂O και αντίστροφα. Συνολικά πρέπει να γίνουν δύο υπολογισμοί στην βελτιστοποιημένη MP2/6-31+G* γεωμετρία του C₂H₄-H₂O. Οι ενέργειες των επιμέρους μορίων έχουν υπολογιστεί στις ενότητες 3.7.2.1 και 3.7.2.2 και είναι:

$$E_{C_2H_4}[CASSCF(2,2)] = -78.06313015 \text{ a.u.}$$

$$E_{H_2O}[RHF] = -76.01653735 \text{ a.u.}$$

Ο υπολογισμός της ενέργειας του H₂O με συναρτήσεις “ghost” στο C₂H₄ είναι τυπικός RHF/6-31+G* και το αποτέλεσμα είναι:



$$E_{\text{H}_2\text{O}}^{\text{GHOST}}[\text{RHF}] = -76.0169464354 \text{ a.u.}$$

Για τον υπολογισμό της ενέργειας CASSCF(2,2)/6-31+G* του C_2H_4 με συναρτήσεις "ghost" στο H_2O χρειάζεται ιδιαίτερη προσοχή διότι πρέπει να μεταβληθεί ο ενεργός χώρος του διμερούς $\text{C}_2\text{H}_4\text{-H}_2\text{O}$ ώστε τα "ghost" μοριακά τροχιακά του H_2O να μην παρεμβάλλονται με τα core και ενεργά μοριακά τροχιακά του C_2H_4 .

Το σύνολο βάσης 6-31+G* του διμερούς $\text{C}_2\text{H}_4\text{-H}_2\text{O}$ αποτελείται από 69 μοριακά τροχιακά τα οποία χρησιμοποιούνται με κατάλληλη αναδιάταξη (εικόνα 3.3) για την εφαρμογή της μεθόδου CASSCF(2,2).

Από τα 69 μοριακά τροχιακά 46 ανήκουν στο C_2H_4 και 23 ανήκουν στο H_2O . Η χρήση συναρτήσεων "ghost" στα 23 μοριακά του H_2O σημαίνει ότι ο υπολογισμός της ολικής ηλεκτρονικής ενέργειας του C_2H_4 γίνεται μόνο για τα 6 άτομα του C_2H_4 αλλά με ολόκληρη την βάση των $46+23=69$ μοριακών τροχιακών του διμερούς.

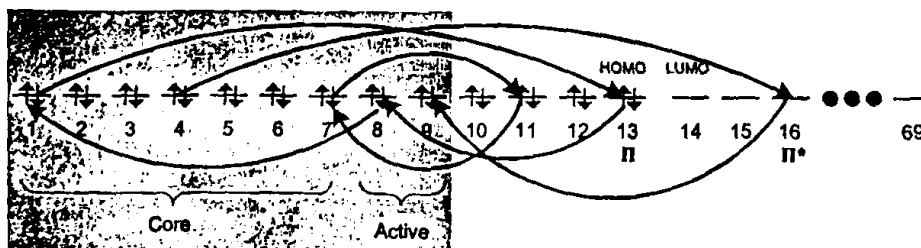
Οπότε ο core και ενεργός χώρος πρέπει να αποτελείται μόνο από μοριακά τροχιακά του C_2H_4 και συνολική βάση 6-31+G* από τα συνολικά 69 μοριακά τροχιακά.

Αυτό σημαίνει ότι για τον core και ενεργό χώρο θα χρησιμοποιηθούν αρχικά μοριακά τροχιακά σαν να υπήρχε το C_2H_4 μόνο του. Από τα αρχικά 69 μοριακά τροχιακά πρέπει αυτά που ανήκουν στο H_2O να βγουν εκτός core και ενεργού χώρου.

Όπως φαίνεται στην εικόνα 3.3 τα 14 μοριακά τροχιακά που χρησιμοποιούνται για τον υπολογισμό CASSCF(2,2)/6-31+G* του διμερούς αποτελούνται από 9 μοριακά τροχιακά του C_2H_4 και 5 μοριακά τροχιακά του H_2O .

Ο CASSCF(2,2)/6-31+G* υπολογισμός του C_2H_4 με "ghost" συναρτήσεις στο H_2O πρέπει να γίνει μόνο με τα 9 μοριακά τροχιακά του C_2H_4 όπου 7 θα είναι core και 2 ενεργά. Στην εικόνα 3.5 φαίνονται οι κατάλληλες αναδιατάξεις που πρέπει να γίνουν στα μοριακά τροχιακά ώστε ο core και ο ενεργός χώρος να αποτελούνται μόνο από τα μοριακά τροχιακά του C_2H_4 .

Εικόνα 3.5



Η αναδιάταξη των μοριακών τροχιακών της εικόνας 3.5 στο αρχείο εισόδου του GAMESS-US γίνεται όπως φαίνεται στην συνέχεια.

```

$guess  guess=MOREAD NORB=69 NORDER=1
        IORDER(1)=8,2,3,5,6,9,11,13,16,10,7,12,1,14,15,4
$end

```

Όλα τα μοριακά τροχιακά του H₂O τα οποία έχουν επιβεβαιωθεί με οπτικοποίηση εξέρχονται από τον core και ενεργό χώρο.

Το αποτέλεσμα του υπολογισμού CASSCF(2,2)/6-31+G* του C₂H₄ με “ghost” συναρτήσεις στο H₂O είναι:

$$E_{\text{C}_2\text{H}_4}^{\text{GHOST}}[\text{CASSCF}(2,2)] = -78.0632149539 \text{ a.u.}$$

Έχοντας ολοκληρωθεί οι επιμέρους υπολογισμοί το BSSE είναι:

$$\begin{aligned} \delta^{\text{BSSE}} &= E_{\text{C}_2\text{H}_4}^{\text{GHOST}}[\text{CASSCF}(2,2)] - E_{\text{C}_2\text{H}_4}[\text{CASSCF}(2,2)] + E_{\text{H}_2\text{O}}^{\text{GHOST}}[\text{RHF}] - E_{\text{H}_2\text{O}}[\text{RHF}] = \\ &= -78.0632149539 + 78.0631301512 - 76.0169464354 + 76.0165373528 = \\ &= -.0004938853 \text{ a.u.} \end{aligned}$$

το οποίο ισοδυναμεί με:

$$\delta^{\text{BSSE}} = -0.30 \text{ Kcal/mol}$$

Στην ενότητα 3.7.2.2 υπολογίστηκε η ενέργεια διαμοριακής αλληλεπίδρασης ίση με -0.99 Kcal/mol . Αν η ενέργεια αυτή διορθωθεί ως προς την ποσότητα BSSE έχουμε την CP-διορθωμένη ενέργεια διαμοριακής αλληλεπίδρασης του διμερούς C₂H₄-H₂O ίση με:

$$\Delta E_{\text{C}_2\text{H}_4-\text{H}_2\text{O}}^{\text{CP}} = -0.69 \text{ Kcal/mol}$$

Αν και δεν μπορεί να γίνει ακόμα ποιοτική αξιολόγηση των αποτελεσμάτων, ποσοτικά η μέθοδος CASSCF(2,2)/6-31+G* επιτυγχάνει στον υπολογισμό της CP-διορθωμένης ενέργειας διαμοριακής αλληλεπίδρασης.

3.8.2 Υπολογισμός του BSSE με την μέθοδο CASSCF(10,7)/6-31+G*

Στην μέθοδο CASSCF(10,7)/6-31+G* θα υπολογιστεί η ενέργεια του C₂H₄ με συναρτήσεις “ghost” στο μόριο του H₂O και αντίστροφα. Συνολικά πρέπει να γίνουν δύο υπολογισμοί στην βελτιστοποιημένη MP2/6-31+G* γεωμετρία του C₂H₄-H₂O. Οι ενέργειες των επιμέρους μορίων έχουν υπολογιστεί στις ενότητες 3.7.3.1 και 3.7.3.2 και είναι:

$$E_{\text{C}_2\text{H}_4}[\text{CASSCF}(2,2)] = -78.06313015 \text{ a.u.}$$

$$E_{\text{H}_2\text{O}}[\text{CASSCF}(8,5)] = -76.0390875 \text{ a.u.}$$



Ο υπολογισμός της ενέργειας CASSCF(2,2)/6-31+G* του C_2H_4 με συναρτήσεις "ghost" στο H_2O γίνεται ακριβώς όπως περιγράφεται στην προηγούμενη υποενότητα και το αποτέλεσμα είναι:

$$E_{C_2H_4}^{GHOST} [CASSCF(2,2)] = -78.0632149539 \text{ a.u.}$$

Ο υπόλογισμός της ενέργειας του H_2O με συναρτήσεις "ghost" στο C_2H_4 γίνεται στην θεωρία CASSCF(8,5)/6-31+G* και πρέπει να μεταβληθεί ο ενεργός χώρος του διμερούς $C_2H_4-H_2O$ ώστε τα "ghost" μοριακά τροχιακά του C_2H_4 να μην παρεμβάλλονται με τα core και ενεργά μοριακά τροχιακά του H_2O .

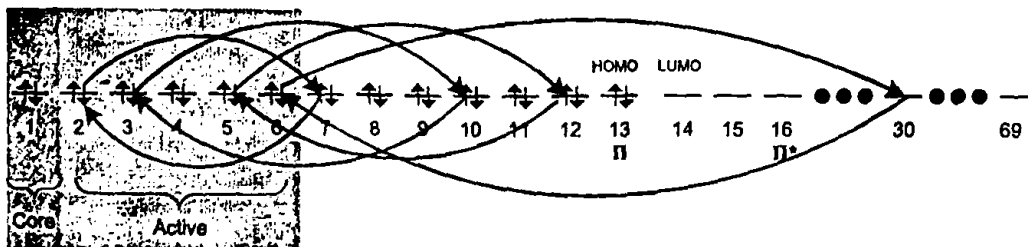
Το σύνολο βάσης 6-31+G* του διμερούς $C_2H_4-H_2O$ αποτελείται από 69 μοριακά τροχιακά τα οποία χρησιμοποιούνται με κατάλληλη αναδιάταξη (εικόνα 3.4) για την εφαρμογή της μεθόδου CASSCF(10,7).

Από τα 69 μοριακά τροχιακά 46 ανήκουν στο C_2H_4 και 23 ανήκουν στο H_2O . Η χρήση συναρτήσεων "ghost" στα 46 μοριακά τροχιακά του C_2H_4 σημαίνει ότι ο υπολογισμός της ολικής ηλεκτρονικής ενέργειας του H_2O γίνεται μόνο για τα 3 άτομα του H_2O αλλά με ολόκληρη την βάση των $46+23=69$ μοριακών τροχιακών του διμερούς.

Οπότε ο core και ενεργός χώρος πρέπει να αποτελείται μόνο από μοριακά τροχιακά του H_2O και η συνολική βάση 6-31+G* από τα συνολικά 69 μοριακά τροχιακά.

Αυτό σημαίνει ότι για τον core και ενεργό χώρο θα χρησιμοποιηθούν αρχικά μοριακά τροχιακά σαν να υπήρχε το μόριο του H_2O μόνο του. Από τα συνολικά 69 μοριακά τροχιακά αυτά που ανήκουν στο C_2H_4 πρέπει να βγουν εκτός core και ενεργού χώρου. Όπως φαίνεται στην εικόνα 3.4 τα 15 μοριακά τροχιακά που χρησιμοποιούνται για τον υπολογισμό CASSCF(10,7)/6-31+G* του διμερούς αποτελούνται από 9 μοριακά τροχιακά του C_2H_4 και 6 μοριακά τροχιακά του H_2O .

Εικόνα 3.6



Ο CASSCF(10,7)/6-31+G* υπολογισμός του H_2O με "ghost" συναρτήσεις στο C_2H_4 πρέπει να γίνει μόνο με τα 6 μοριακά τροχιακά του H_2O όπου 1 θα είναι core και 5



ενεργά. Στην εικόνα 3.6 φαίνονται οι κατάλληλες αναδιατάξεις που πρέπει να γίνουν στα μοριακά τροχιακά ώστε ο core και ο ενεργός χώρος να αποτελούνται μόνο από τα μοριακά τροχιακά του H₂O.

Η αναδιάταξη των μοριακών τροχιακών της εικόνας 3.6 στο αρχείο εισόδου του GAMESS-US γίνεται όπως φαίνεται στην συνέχεια:

```
$guess guess=MOREAD NORB=69 NORDER=1
      IORDER(2)=7,10,4,12,30,2,8,9,3,11,5
      IORDER(30)=6
$end
```

Όλα τα μοριακά τροχιακά του C₂H₄ τα οποία έχουν επιβεβαιωθεί με οπτικοποίηση εξέρχονται από τον core και ενεργό χώρο.

Το αποτέλεσμα του υπολογισμού CASSCF(8,5)/6-31+G* του H₂O με “ghost” συναρτήσεις στο C₂H₄ είναι:

$$E_{\text{H}_2\text{O}}^{\text{GHOST}}[\text{CASSCF}(8,5)] = -76.03959923 \text{ a.u.}$$

Έχοντας ολοκληρωθεί οι επιμέρους υπολογισμοί το BSSE είναι:

$$\begin{aligned} \delta^{\text{BSSE}} &= E_{\text{C}_2\text{H}_4}^{\text{GHOST}}[\text{CASSCF}(2,2)] - E_{\text{C}_2\text{H}_4}[\text{CASSCF}(2,2)] + \\ &+ E_{\text{H}_2\text{O}}^{\text{GHOST}}[\text{CASSCF}(8,5)] - E_{\text{H}_2\text{O}}[\text{CASSCF}(8,5)] = \\ &- 78.0632149539 + 78.0631301512 - 76.03959923 + 76.0390875 = \\ &- 0.0005965327 \text{ a.u.} \end{aligned}$$

το οποίο ισοδυναμεί με:

$$\delta^{\text{BSSE}} = -0.37 \text{ Kcal/mol}$$

Στην ενότητα 3.7.3.2 υπολογίστηκε η ενέργεια διαμοριακής αλληλεπίδρασης ίση με -1.03 Kcal/mol . Αν η ενέργεια αυτή διορθωθεί ως προς την ποσότητα BSSE έχουμε την CP-διορθωμένη ενέργεια διαμοριακής αλληλεπίδρασης του διμερούς C₂H₄-H₂O ίση με:

$$\Delta E_{\text{C}_2\text{H}_4-\text{H}_2\text{O}}^{\text{CP}} = -0.66 \text{ Kcal/mol}$$

Αν και δεν μπορεί να γίνει ακόμα ποιοτική αξιολόγηση των αποτελεσμάτων, ποσοτικά η μέθοδος CASSCF(10,7)/6-31+G* επιτυγχάνει στον υπολογισμό της CP-διορθωμένης ενέργειας διαμοριακής αλληλεπίδρασης.



3.9. Υπολογισμός του BSSE με την μέθοδο NEVPT2

Στην συνέχεια θα προσθέσουμε δυναμική ηλεκτρονική συσχέτιση εφαρμόζοντας την μέθοδο NEVPT2 στους υπολογισμούς της ενότητας 3.8. Η επιλογή και των δύο ενεργών χώρων θα περιέχει τα μοριακά τροχιακά που ήδη έχουν εντοπιστεί. Στόχος είναι η βελτίωση των αποτελεσμάτων για τον υπολογισμό της διαμοριακής ενέργειας αλληλεπίδρασης. Όλοι οι υπολογισμοί έγιναν με τα πακέτα GAMESS-US και Dalton 2.0. Τα βελτιστοποιημένα CASSCF μοριακά τροχιακά χρησιμοποιήθηκαν ως αρχικά στο πακέτο Dalton ώστε να αποφευχθούν διπλές αναδιατάξεις, όπως φαίνονται στις εικόνες 3.5 και 3.6.

Τα αποτελέσματα των υπολογισμών φαίνονται στον πίνακα 3.1. Η μέθοδος NEVPT2 είναι σε πλήρη συμφωνία με την μέθοδο MP2 στην περίπτωση του υπολογισμού που αφορά το μονομερές H₂O στην περίπτωση του ενεργού χώρου NEVPT2(2,2)/6-31+G*. Η δύο κατηγορίες SC και PC της μεθόδου NEVPT2 επικυρώνουν την size-consistency για όλα τα ενεργειακά αποτελέσματα.

Πίνακας 3.1: Τα ενεργειακά αποτελέσματα από τους υπολογισμούς NEVPT2(10,7)/6-31+G* και NEVPT2(2,2)/6-31+G* στην MP2/6-31+G* βελτιστοποιημένη μοριακή γεωμετρία.

Active Space	Theory	Eth-Water Complex	Monomers	Ghost Basis		Interaction Kcal/mol		δ^{BSSE} Kcal/mol
AB ⁽¹⁾		E _{AB} (AB)	E _A (A)+E _B (B)	E _A (AB)	E _B (AB)	$\Delta E^{(2)}$	$\Delta E_{CP}^{(3)}$	
10,7	CASSCF	-154.10386090	-154.10221763	-78.06321495	-76.03959923	-1.03	-0.66	-0.37
	NEVPT2 (SC)	-154.50146957	-154.49700326	-78.30623862	-76.19286636	-2.80	-1.48	-1.32
	NEVPT2 (PC)	-154.50440783	-154.49966987	-78.30630703	-76.19541629	-2.97	-1.68	-1.29
2,2	CASSCF	-154.08124464	-154.07966753	-78.06321495	-76.01694641	-0.99	-0.68	-0.31
	NEVPT2 (SC)	-154.52204591	-154.51746019	-78.30623862	-76.21323151	-2.88	-1.61	-1.27
	NEVPT2 (PC)	-154.52211224	-154.51753196	-78.30630703	-76.21323151	-2.87	-1.61	-1.26

⁽¹⁾A= C₂H₄, B=H₂O, AB= C₂H₄-H₂O. ⁽²⁾Η μη διορθωμένη διαμοριακή ενέργεια. ⁽³⁾Η CP-διορθωμένη διαμοριακή ενέργεια αλληλεπίδρασης.

Η μελέτη του διμερούς C₂H₄-H₂O επεκτάθηκε με υπολογισμούς NEVPT2 χρησιμοποιώντας μεγαλύτερα σύνολα βάσεων στην αντίστοιχη βελτιστοποιημένη MP2 μοριακή γεωμετρία. Συγκεκριμένα το διμερές βελτιστοποιήθηκε στις εξής θεωρίες:



- G1. MP2/6-311++G**
 G2. MP2/aug-cc-pVDZ
 G3. MP2/ANO-1
 G4. MRMP2(10,7)/6-311++G**

Το σύνολο βάσεων ANO-1 της ομάδας του Roos εφαρμόστηκε χρησιμοποιώντας την αντιστοιχία [3s2p] για τα άτομα του H και [4s3p2d] για τα άτομα C και O¹⁶⁹.

Το σύνολο βάσης 6-311++G** χρησιμοποιήθηκε διότι συμπεριφέρεται ποιοτικά καλύτερα ως σύνολο τριπλού-ζήτα σε σύγκριση με το σύνολο βάσεων aug-cc-pVDZ το οποίο είναι τύπου διπλού-ζήτα¹⁷⁰.

3.9.1. Επέκταση της διαδικασίας σε άλλες βάσεις

Σε προηγούμενες θεωρητικές και πειραματικές μελέτες¹⁷¹ η διαμοριακή ενέργεια αλληλεπίδρασης βρέθηκε μεταξύ 1.50–2.0 kcal/mol. Σύμφωνα με τον πίνακα 3.1, οι CP-διορθωμένες NEVPT2/6-31+G* διαμοριακές ενέργειες αλληλεπίδρασης ΔE_{CP} βρίσκονται σε πολύ καλή συμφωνία με την βιβλιογραφία.

Ο μεγαλύτερος χώρος της προσέγγισης PC-NEVPT2 παρουσιάζει καλύτερα αποτελέσματα στο σύνολο βάσης 6-31+G*. Όπως είδαμε και στις προηγούμενες υποενότητες, οι υπολογισμοί CASSCF(2,2)/6-31+G* και CASSCF(10,7)/6-31+G* απέτυχαν να αποδώσουν μία καλή προσέγγιση για την CP-διορθωμένη διαμοριακή ενέργεια αλληλεπίδρασης του διμερούς.

Πίνακας 3.2: Η μη-διορθωμένη, η CP-διορθωμένη και το BSSE για τους υπολογισμούς NEVPT2 στο διμερές C₂H₄-H₂O στις γεωμετρίες G1,G2,G3 και G4. Όλες οι τιμές είναι σε Kcal/mol.

Geometry NEVPT2 Basis	G1 6-311++G**			G2 aug-cc-pVDZ			G3 ANO-1			G4 6-311++G**		
	CAS ⁽¹⁾	SC ⁽²⁾	PC ⁽³⁾	CAS	SC	PC	CAS	SC	PC	CAS	SC	PC
(10,7) Active Space												
ΔE	-0.49	-1.63	-1.85	-0.30	-2.45	-2.75	-0.09	-3.21	-3.53	-0.63	-1.71	-1.91
ΔE _{CP}	-0.32	-0.79	-0.99	-0.03	-1.01	-1.25	0.01	-1.04	-1.29	-0.48	-0.90	-1.08
BSSE	-0.16	-0.84	-0.86	-0.27	-1.44	-1.49	-0.10	-2.17	-2.24	-0.15	-0.81	-0.83
(2,2) Active Space												
ΔE	-0.82	-2.56	-2.56	-0.71	-3.56	-3.55	-0.44	-4.38	-4.38	-0.93	-2.57	-2.57
ΔE _{CP}	-0.65	-1.65	-1.65	-0.45	-2.01	-2.01	-0.27	-1.97	-1.97	-0.77	-1.70	-1.70
BSSE	-0.17	-0.91	-0.91	-0.27	-1.55	-1.54	-0.18	-2.41	-2.41	-0.15	-0.87	-0.87

⁽¹⁾ Υπολογισμοί CASSCF calculation. ⁽²⁾ Προσέγγιση NEVPT2-SC. ⁽³⁾ Προσέγγιση NEVPT2-PC.



Στο πίνακα 3.2 παρουσιάζονται τα αποτελέσματα από τους υπολογισμούς NEVPT2, οι οποίοι έγιναν στις βελτιστοποιημένες γεωμετρίες G1, G2, G3 και G4.

Σε όλες τις περιπτώσεις η προσέγγιση PC-NEVPT2 δίνει πιο ακριβή και αξιόπιστα αποτελέσματα σε σύγκριση με την προσέγγιση SC-NEVPT2. Τα ενεργειακά αποτελέσματα ΔE_{CP} στον ενεργό χώρο (2,2), βρίσκονται σε καλή συμφωνία με προηγούμενες πειραματικές και θεωρητικές προσπάθειες. Το φαινόμενο της ελάττωσης του BSSE καθώς το σύνολο βάσης αυξάνεται, παρατηρείται σε όλες τις περιπτώσεις. Τα ισοδύναμα αποτελέσματα των βάσεων aug-cc-pVDZ και ANO-1 δικαιολογούνται από την διπλού-ζήτα ποιότητα των δύο βάσεων.

Οι αρκετά διαφορετικές G1 και G4 βελτιστοποιημένες γεωμετρίες παρέχουν παρόμοια ποσοτικά αποτελέσματα στην μέθοδο NEVPT2/6-311++G**. Επίσης, η πειραματική διαπίστωση της πολύ ασθενούς διαμοριακής αλληλεπίδρασης επιβεβαιώνεται από τους NEVPT2/6-311++G** υπολογισμούς.

Για να ελεγχθεί η ακρίβεια της μεθοδολογίας των ενεργειακών αποτελεσμάτων NEVPT2 έγιναν επιπλέον υπολογισμοί με την θεωρία CCSD(T) στις βελτιστοποιημένες γεωμετρίες και βάσεις, ανάλογα με τον πίνακα 3.2.

Πίνακας 3.3: Η μη-διορθωμένη, η CP-διορθωμένη και το BSSE για τους υπολογισμούς CCSD(T) για το διμερές $C_2H_4-H_2O$ στις γεωμετρίες G1, G2, G3 και G4. Όλες οι τιμές είναι σε Kcal/mol.

Geometry CCSD(T) Basis	G1 6-311++G**	G2 aug-cc-pVDZ	G3 ANO-1	G4 6-311++G**
ΔE	-2.63	-3.67	-4.49	-2.64
ΔE_{CP}	-1.63	-2.01	-2.01	-1.68
BSSE	-1.00	-1.65	-2.48	-0.96

Στον πίνακα 3.3 απεικονίζονται τα αποτελέσματα του υπολογισμού CCSD(T) ως στοιχείο σύγκρισης των υπολογισμών NEVPT2. Η CP-διορθωμένη διαμοριακή ενέργεια αλληλεπίδρασης εμφανίζει ακριβώς τις ίδιες τιμές σε σύγκριση με τους υπολογισμούς NEVPT2 στον ενεργό χώρο (2,2).

3.9.2. Επέκταση της μεθόδου για την μελέτη του διμερούς $C_6H_6-H_2O$

Η υπολογιστική διαδικασία που αναπτύχθηκε σε αυτό το κεφάλαιο, εφαρμόστηκε επιπλέον στην μελέτη της διαμοριακής αλληλεπίδρασης του διμερούς $C_6H_6-H_2O$. Επιλέχθηκε το βέλτιστο σημείο της υπερεπιφάνειας δυναμικής ενέργειας, όπως αυτό υπολογίστηκε στο κεφάλαιο 2. Η βέλτιστη γεωμετρική¹⁷² διεύθυνση του διμερούς



απαιτεί το άτομο του O να βρίσκεται στον C_6 άξονα συμμετρίας του βενζολίου και τα άτομα του H να έχουν κατεύθυνση προς το επίπεδο του C_6H_6 .

Πίνακας 3.4: Τα ενεργειακά αποτελέσματα NEVPT2(6,6)/6-31G** του διμερούς $C_6H_6-H_2O$ στην MP2/6-31G** βελτιστοποιημένη γεωμετρία (βλ. κεφάλαιο 2).

Active Space	Theory	Eth-Water Complex	Monomers	Ghost Basis		Interaction Kcal/mol		δ^{BSSE} Kcal/mol
AB ⁽¹⁾		$E_{AB}(AB)$	$E_A(A)+E_B(B)$	$E_A(AB)$	$E_B(AB)$	$\Delta E^{(2)}$	$\Delta E_{CP}^{(3)}$	
6,6	CASSCF	-306.81307843	-306.80952443	-230.78764062	-76.02319278	-2.23	-1.41	-0.82
	NEVPT2 (SC)	-307.76826517	-307.76238099	-231.54143727	-76.22333726	-3.69	-2.19	-1.50
	NEVPT2 (PC)	-307.76888266	-307.76298235	-231.54205706	-76.22333726	-3.70	-2.19	-1.51

⁽¹⁾A= C_6H_6 , B= H_2O , AB= $C_6H_6-H_2O$. ⁽²⁾Η μη διορθωμένη διαμοριακή ενέργεια. ⁽³⁾Η CP-διορθωμένη διαμοριακή ενέργεια αλληλεπίδρασης.

Η μέθοδος NEVPT2 εφαρμόστηκε στο διμερές σύμφωνα με την μεθοδολογία επιλογής του ενεργού χώρου που αναπτύχθηκε. Ο ενεργός χώρος αποτελείται από 6 π και π^* μοριακά τροχιακά του C_6H_6 και 23 core μοριακά τροχιακά. Τα αρχικά μοριακά τροχιακά εντοπίστηκαν στο επίπεδο θεωρίας RHF/6-31G**. Τα αποτελέσματα παρουσιάζονται στον πίνακα 3.4 και είναι σε πολύ καλή συμφωνία με τα πειραματικά αποτελέσματα που κυμαίνονται μεταξύ 1.4–2.8 kcal/mol.

Οι υπολογισμοί NEVPT2(6,6) στο διμερές $C_6H_6-H_2O$ είναι μια ακριβής και αξιόπιστη επέκταση των υπολογισμών NEVPT2(2,2) που έγιναν για το διμερές $C_2H_4-H_2O$. Η μόνη διαφορά της μεθοδολογίας που αναπτύχθηκε αφορά την επιλογή των μοριακών τροχιακών.

3.10. Συμπεράσματα

Σε αυτό το κεφάλαιο αναπτύχθηκε μία διαδικασία μοριακής προσομοίωσης η οποία μπορεί να εφαρμοστεί στην μελέτη της CP-διορθωμένης OH..π διαμοριακής ενέργειας αλληλεπίδρασης σε ετεροδιμερή μοριακά συστήματα.

Η ιδιότητα της size-consistency που παρέχει η μέθοδος NEVPT2 αποτελεί ένα πολύ μεγάλο πλεονέκτημα για την εφαρμογή της μεθόδου ώστε να υπολογιστούν με μεθόδους υψηλής ακρίβειας και αξιοπιστίας οι διαμοριακές αλληλεπιδράσεις ετεροδιμερών μοριακών συστημάτων.



Η προσέγγιση PC-NEVPT2 στους δύο ενεργούς χώρους που έγιναν οι υπολογισμοί έδωσε καλύτερα ποσοτικά αποτελέσματα, όπως αναμενόταν, σε σύγκριση με την SC-NEVPT2.

Τα πειραματικά αποτελέσματα, για όλες τις περιπτώσεις που εφαρμόστηκε η θεωρία NEVPT2, συγκρινόμενα με τα αποτελέσματα της νέας υπολογιστικής διαδικασίας είναι σε πολύ καλή συμφωνία.

Η μεθοδολογία αναπτύχθηκε για το μοριακό σύστημα $C_2H_4-H_2O$ και στη συνέχεια επεκτάθηκε στο μεγαλύτερο $C_6H_6-H_2O$ όπου έδωσε αξιόπιστα αποτελέσματα και ελέγχθηκε ποσοτικά.

Τελικά ο πολύ-οριζουσιακός χαρακτήρας των CP-διορθωμένων OH..π διαμοριακών αλληλεπιδράσεων, που παίζουν πολύ σημαντικό ρόλο στην διευθέτηση μεγάλων μοριακών συστημάτων βιολογικής σημασίας, μπορεί να εμπλουτιστεί με την εφαρμογή νέων αξιόπιστων μεθοδολογιών που στοχεύουν σε αυτοματοποιημένους και υψηλής ακρίβειας υπολογισμούς μοριακής προσομοίωσης.



ΚΕΦΑΛΑΙΟ 4

TINGAM: Δημιουργία ενός ενοποιημένου υπολογιστικού συστήματος με εφαρμογές σε μοριακά συστήματα βιολογικής σημασίας.



4.1. Εισαγωγή

Η *ab-initio* μοριακή δυναμική είναι ένα από τα πιο ισχυρά εργαλεία για την μελέτη μοριακών συστημάτων. Πρόσφατα¹⁷³ η ενοποίηση ενός γενικού πακέτου μοριακής δυναμικής με το GAUSSIAN χρησιμοποιήθηκε για την μελέτη της δυναμικής εξέλιξης καθώς και των δονητικών ιδιοτήτων ενός μορίου που βρίσκεται σε διεγερμένη κατάσταση. Σε μία άλλη εργασία¹⁷⁴ έγινε υπολογιστική προσομοίωση της διεγερμένης PES της προπανάλης με την μέθοδο AIMS (*ab-initio multiple spawning*) και χρησιμοποιήθηκε non-BO προσέγγιση με την CASSCF(5,5) μέθοδο για τον υπολογισμό της ηλεκτρονικής ενέργειας. Η μέθοδος AIMS λύνει την εξίσωση Schrodinger τόσο για τα ηλεκτρόνια όσο και για τους πυρήνες ταυτόχρονα, επιτρέποντας μια ευέλικτη περιγραφή του συστήματος.

Στο κεφάλαιο αυτό γίνεται η διασύνδεση των πακέτων GAMESS-US και TINKER με σκοπό την επίτευξη *ab-initio* προσομοιώσεων μοριακής δυναμικής. Η διασύνδεση αυτή πραγματοποιείται με την ανάπτυξη μίας βιβλιοθήκης σε γλώσσα προγραμματισμού ANSI C, η οποία παρέχει την διαλειτουργικότητα μεταξύ των δύο πακέτων. Για την υλοποίηση του ενοποιημένου συστήματος αναπτύσσονται ειδικές ρουτίνες σε Fortran 77 και ενσωματώνονται στο πακέτο TINKER. Η βιβλιοθήκη που δημιουργήθηκε αποτελεί μία επέκταση του συστήματος MERGAM, το οποίο παρουσιάστηκε στο κεφάλαιο 2 και ονομάζεται TINGAM. Χρησιμοποιώντας την τεχνολογία ανάπτυξης ρουτινών που παρέχει το TINKER και διατηρώντας όλα τα χαρακτηριστικά της κλασικής μοριακής δυναμικής, αναπτύχθηκε μία βιβλιοθήκη διασύνδεσης του πακέτου TINKER με το πακέτο GAMESS-US. Με το νέο ενοποιημένο σύστημα είναι δυνατός ο υπολογισμός της τροχιάς (*trajectory*) ενός μοριακού συστήματος, του οποίου η δυναμική ενέργεια υπολογίζεται με *ab-initio* ή ημι-εμπειρικές μεθόδους μέσω του πακέτου GAMESS-US.

Το ενοποιημένο *ab-initio* σύστημα μοριακής δυναμικής που αναπτύχθηκε, ελέγχθηκε με εφαρμογές στο διπεπτίδιο της αλανίνης και στην κυκλό-εξαγλυκίνη με την ημιεμπειρική μέθοδο AM1 και το μοντέλο επιδιάλυτωσης SM5.42. Επίσης, εφαρμόστηκε για να διερευνηθούν τα διαμορφωσιακά χαρακτηριστικά του πεπτιδικού μοντέλου των σερινοπρωτεασών.

Στις επόμενες ενότητες παρουσιάζεται η θεωρία των προσομοιώσεων της κλασικής μοριακής δυναμικής καθώς και οι αρχές των *ab-initio* προσομοιώσεων. Ακολουθεί η περιγραφή της πλήρους ανάπτυξη του ενοποιημένου συστήματος και οι εφαρμογές.



4.2. Μοριακή Δυναμική

Η Μοριακή Δυναμική (ΜΔ) λύνει τις κλασσικές εξισώσεις κίνησης¹⁷⁵ για ένα σύστημα N ατόμων που αλληλεπιδρούν με τη δυναμική ενέργεια ενός πεδίου δυνάμεων. Η λύση της εξίσωσης κίνησης ενός μοριακού συστήματος αφορά την χρονική εξέλιξη των κινήσεων των ατόμων του. Γνωρίζοντας λοιπόν τον χρόνο εξέλιξης μπορούν να μελετηθούν χρονικά εξαρτημένες ιδιότητες του μορίου, όπως είναι η διασπορά, διάχυση κ.λ.π. Η μοριακή δυναμική επίσης δεν περιορίζεται σε αρμονική κίνηση γύρω από ένα απλό ελάχιστο. Έτσι επιτρέπει στα μόρια να διασταυρωθούν με ενεργειακά μέγιστα ώστε να μπορέσουν να εξερευνηθούν και άλλες περιοχές του διαμορφωσιακού τους χώρου.

Η μελέτη της εξέλιξης ενός μοριακού συστήματος στον χρόνο είναι το πλεονέκτημα της μοριακής δυναμικής υπό τον περιορισμό ότι ο συνολικός χρόνος που πρακτικά μπορεί να προσομοιωθεί είναι της τάξεως 10^{-12} έως 10^{-9} seconds (picoseconds έως nanoseconds). Η προσομοίωση 1 ps με δειγματοληψία διαμορφώσεων ανά 1 fs δημιουργεί 1000 διαμορφώσεις. Αντίστοιχα, η προσομοίωση 1 ns με το ίδιο βήμα δειγματοληψίας, δημιουργεί ένα εκατομμύριο διαμορφώσεις.

4.2.1. Εφαρμογές της Μοριακής Δυναμικής

Οι προσομοιώσεις Μοριακής Δυναμικής έχουν εφαρμογή στη μελέτη μεγάλων μοριακών συστημάτων όπως: βιολογικά μόρια, πολυμερή, υδατικά διαλύματα ή ακόμα και ενώσεις σε αέρια φάση. Οι σημαντικότερες εφαρμογές της μοριακής δυναμικής¹⁷⁶ επικεντρώνονται στα εξής:

- Έρευνα του διαμορφωσιακού χώρου.
- Παραγωγή στατιστικών συνόλων.
- Μελέτη των μοριακών κινήσεων.

Για τον υπολογισμό θερμοδυναμικών ιδιοτήτων παράγονται στατιστικά σύνολα όπως τα μικροκανονικά ή τα κανονικά.. Σε αυτές τις μελέτες είναι σημαντικό το ότι ο υπολογισμός περνά από ποικίλα στάδια των διαμορφώσεων έχοντας τη σωστή στατιστική συχνότητα. Μπορούν να χρησιμοποιηθούν οι μελέτες των μοριακών κινήσεων, για να βρεθούν ιδιότητες όπως π.χ οι συντελεστές διάχυσης κ.τ.λ.

Η μελέτη των δυναμικών κινήσεων των μορίων παρουσιάζει μεγάλο ενδιαφέρον για την χημεία και βιοχημεία. Παρόλο που η σύγχρονη κρυσταλλογραφία έχει εξασφαλίσει επαρκείς μεθόδους για τη στατική δομή μικρών και μεγάλων μορίων,



υπάρχει πάντα ενδιαφέρον για τις ενδομοριακές συγκρούσεις και τις μεταβολές των διαμορφώσεων.

Μία άλλη προσέγγιση για την προσομοίωση της μοριακής κίνησης είναι η μέθοδος Monte Carlo η οποία βασίζεται στο γεγονός ότι τα μόρια του συστήματος κινούνται τυχαία σύμφωνα με στατιστικούς κανόνες. Ενώ αυτή η μέθοδος μπορεί να παίρνει δείγματα του διαμορφωσιακού χώρου έτσι ώστε να παράγει μικρής σημασίας στατιστικά σύνολα, δεν παρέχει δυναμικές πληροφορίες του συστήματος.

4.2.2. Εξισώσεις Μοριακής Δυναμικής

Η μοριακή δυναμική επιλύει τη εξίσωση κίνησης του Newton,

$$F_i(t) = m_i a_i(t) \quad (4.1)$$

όπου:

- $F_i(t)$ είναι η δύναμη,
- m_i είναι η μάζα,
- $a_i(t)$ είναι η επιτάχυνση του ατόμου i .

Η δύναμη πάνω στο άτομο i μπορεί να υπολογιστεί απ' ευθείας από την παράγωγο της δυναμικής ενέργειας V συναρτήσεως των ατομικών συντεταγμένων R_i . Αυτό συνεπάγεται ότι η ενέργεια ενός μοριακού συστήματος εκφράζεται με την διαφορική μορφή:

$$-\frac{\partial V}{\partial R_i} = m_i \frac{\partial^2 R_i}{\partial t^2} \quad (4.2)$$

Η πλήρης γνώση του δυναμικού και των μαζών θα ήταν πιθανό να λύσει τη διαφορική εξίσωση (4.2), για όλες τις θέσεις των ατόμων ενός μορίου στο χρόνο. Εφόσον οι αρχικές συντεταγμένες και ταχύτητες είναι γνωστές τότε μπορούν να υπολογιστούν οι συντεταγμένες και οι ταχύτητες στην επόμενη χρονική στιγμή. Οι συντεταγμένες και οι ταχύτητες των ατόμων για μία ολοκληρωμένη δυναμική πορεία αποτελούν την τροχιά (trajectory) που ακολουθεί το σύστημα.

Η κλασσική μηχανική παράγει περιορισμένες λύσεις της εξίσωσης (4.2), γι' αυτό στην πράξη απαιτούνται αριθμητικές μέθοδοι, για την επίλυση μεγαλύτερων συστημάτων. Για την αριθμητική επίλυση της κίνησης ενός ατόμου αναπτύσσεται μια σειρά Taylor. Εάν η θέση του ατόμου την χρονική στιγμή t είναι $R(t)$ τότε η θέση του μετά από μικρό χρονικό διάστημα Δt είναι:



$$R(t + \Delta t) = R(t) + \frac{dR}{dt} \Delta t + \frac{d^2 R}{dt^2} \frac{\Delta t^2}{2} + \dots \quad (4.3)$$

Η αριθμητική λύση της εξίσωσης κίνησης (4.3) εξαρτάται από δύο παράγοντες.

1. Τη γνώση της θέσης $R(t)$, άρα και της ταχύτητας $\frac{dR}{dt}$ και της επιτάχυνσης $\frac{d^2 R}{dt^2}$
2. Τη γνώση άλλων προσεγγίσεων που συνεισφέρουν αντί των όρων υψηλής τάξης της ανάπτυξης Taylor.

4.2.3. Τα βήματα μιας προσομοίωσης μοριακής δυναμικής

Για την επίτευξη μιας προσομοίωσης μοριακής δυναμικής ακολουθείται η παρακάτω διαδικασία:

- Ορίζονται οι αρχικές συντεταγμένες των ατόμων του μοριακού συστήματος.
- Οι εσωτερικές ταχύτητες ορίζονται τυχαία από την κατανομή Maxwell-Boltzman και ισοδυναμούν με τις θερμοκρασίες για τις οποίες υπάρχει ενδιαφέρον.
- Η επιτάχυνση υπολογίζεται από την εξίσωση (4.2).
- Υπολογίζεται η νέα θέση $R(t+\Delta t)$ από την εξίσωση (4.3).
- Αντικαθίστανται οι αρχικές συντεταγμένες με τις νέες συντεταγμένες.
- Οι νέες ταχύτητες ενημερώνονται, αφαιρώντας από την προηγούμενη τιμή των συντεταγμένων τις νέες συντεταγμένες και διαιρώντας με το βήμα του χρόνου που ακολουθήθηκε.
- Η νέα επιτάχυνση διορθώνεται υπολογίζοντας την κλίση της ταχύτητας με τις καινούργιες συντεταγμένες.

Ένα βήμα (step) έχει ολοκληρωθεί αφού γίνουν τα παραπάνω. Τα βήματα αυτά επαναλαμβάνονται μέχρι το σύστημα να ισορροπήσει.

4.3. Το πακέτο Μοριακής Δυναμικής TINKER

Το TINKER¹⁷⁷ είναι ένα σύστημα που αποτελείται από διάφορα προγράμματα και εργαλεία για υπολογισμούς μοριακής μηχανικής και προσομοιώσεις μοριακής δυναμικής. Είναι ένα ολοκληρωμένο πακέτο γενικής χρήσης για μοριακή μοντελοποίηση ενώ έχει ειδικά χαρακτηριστικά για την μελέτη βιο-πολυμερών. Παρέχει ένα μεγάλο πλήθος από πεδία δυνάμεων (force fields) μοριακής μηχανικής όπως Amber, Allinger MM (MM2-1991 και MM3-2000), OPLS (OPLS-UA, OPLS-AA και OPLS-AA/L) κ.α. Επίσης, αποτελείται από νέους αλγόριθμους για εύρεση



τοπικών και ολικών ελάχιστων καθώς για την εξερεύνηση της υπερεπιφάνειας δυναμικής ενέργειας. Όλα τα προγράμματα και εργαλεία του TINKER είναι γραμμένα σε γλώσσα προγραμματισμού Fortran 77.

Το πακέτο αναπτύχθηκε με σκοπό να προσφέρει μια πλατφόρμα πάνω στην οποία μπορούν να αναπτυχθούν νέες μέθοδοι και αλγόριθμοι στο πεδίο της μοριακής μοντελοποίησης¹⁷⁸. Η καρδιά του πακέτου αποτελείται από ένα σύνολο ρουτίνων οι οποίες επιτρέπουν την διαχείριση όλων των εφαρμογών του. Σε αυτό το πλαίσιο στηρίζεται και η ανάπτυξη που αναφέρεται σε αυτό το κεφάλαιο και αποτελεί μία επέκταση των υπολογισμών της μοριακής δυναμικής χρησιμοποιώντας ab-initio μοριακούς κβαντικούς υπολογισμούς διασυνδέοντας το TINKER με το πακέτο GAMESS.

4.3.1. Το πρόγραμμα μοριακής δυναμικής dynamic του TINKER.

Το πρόγραμμα που παρέχει το TINKER για προσομοιώσεις μοριακής δυναμικής ονομάζεται **dynamic**. Είναι μία εφαρμογή που αναπτύχθηκε με τις ρουτίνες βιβλιοθήκης που παρέχει το TINKER σε Fortran 77. Το πρόγραμμα εκτελεί υπολογισμούς μοριακής δυναμικής χρησιμοποιώντας ως είσοδο ένα αρχείο που περιγράφει τις καρτεσιανές συντεταγμένες των ατόμων του μορίου.

Η τροχιά της μοριακής δυναμικής στην διάρκεια του χρόνου υπολογίζεται είτε με τον αλγόριθμο ολοκλήρωσης modified Beeman¹⁷⁹ είτε με τον velocity Verlet^{180, 181}. Οι αλγόριθμοι ολοκλήρωσης της μοριακής δυναμικής θα αναλυθούν στην συνέχεια. Για την εκτέλεση ενός υπολογισμού ο χρήστης επιλέγει και εισάγει:

- a) Τον αλγόριθμο ολοκλήρωσης.
- b) Την αρχική θερμοκρασία ή πίεση.
- c) Τον αριθμό των βημάτων της μοριακής δυναμικής.
- d) Το χρονικό διάστημα στο οποίο θα εκτελεστούν τα βήματα.
- e) Ένα διάστημα βημάτων στο οποίο θα αποθηκεύονται τα σημεία της τροχιάς.

Εφόσον κατά την διάρκεια υπολογισμού της τροχιάς τα σημεία αποθηκεύονται, υπάρχει η δυνατότητα ένας υπολογισμός να συνεχίσει χρησιμοποιώντας οποιοδήποτε αποθηκευμένο σημείο ως αρχικό. Αυτό είναι χρήσιμο στην περίπτωση που ο υπολογισμός διακοπεί λόγω εξωτερικών παραγόντων. Το αρχείο που καταγράφεται η τροχιά έχει κατάληξη dyn. Η επιλογή του αλγόριθμου ολοκλήρωσης γίνεται με την λέξη κλειδί (keyword) INTEGRATE.



4.3.2. Αλγόριθμοι ολοκλήρωσης

Μία αριθμητική μέθοδος που λύνει μία συνηθισμένη διαφορική εξίσωση όπως η εξίσωση (4.2) είναι η μέθοδος των μερικών διαφορικών. Η γενική ιδέα είναι η ακόλουθη. Δίνοντας τις αρχικές συντεταγμένες και ταχύτητες την χρονική στιγμή t , υπολογίζονται οι θέσεις και οι ταχύτητες σε χρόνο $t+\Delta t$. Το βήμα του χρόνου Δt εξαρτάται από την μέθοδο ολοκλήρωσης καθώς και από το ίδιο το μοριακό σύστημα.

Οι περισσότερο χρησιμοποιούμενες μέθοδοι ολοκλήρωσης των εξισώσεων κίνησης, στη μοριακή δυναμική είναι οι ποικίλες μορφές των αλγορίθμων ολοκλήρωσης του Verlet. Τα πλεονεκτήματα των αλγορίθμων του Verlet σε σχέση με τους υπόλοιπους είναι ότι απαιτεί μόνο ένα υπολογισμό της ενέργειας για κάθε βήμα, απαιτεί επίσης μόνο την ελάχιστη μνήμη και επιτρέπει ένα σχετικά μεγάλο μήκος βήματος χρόνου.

4.3.2.1. Μέθοδος Verlet Leapfrog

Η μέθοδος ολοκλήρωσης που χρησιμοποιείται συνήθως για τη λύση της εξίσωσης (4.3) περιγράφηκε πρώτα από τον Verlet (1967). Σύμφωνα με τη μέθοδο αυτή θεωρείται ότι:

v_{av} είναι η μέση τιμή της ταχύτητας κατά τη διάρκεια του διαστήματος χρόνου t έως $t + \Delta t$.

Τότε η θέση ενός ατόμου στο τέλος του διαστήματος του χρόνου είναι:

$$R(t + \Delta t) = R(t) + v_{av}\Delta t \quad (4.4)$$

Εάν θεωρηθεί ότι η ταχύτητα μεταβάλλεται γραμμικά στο χρόνο Δt τότε η μέση τιμή της ταχύτητας μπορεί να εξισωθεί με τη στιγμιαία ταχύτητα στο χρόνο $\frac{\Delta t}{2}$.

$$v_{av} = v\left(t + \frac{\Delta t}{2}\right) \quad (4.5)$$

Η στιγμιαία ταχύτητα στη συνέχεια μπορεί να υπολογιστεί από τη μέση τιμή επιταχύνσεων a_{av} για τον χρόνο από $\left(t - \frac{\Delta t}{2}\right)$ έως $\left(t + \frac{\Delta t}{2}\right)$.

$$v\left(t + \frac{\Delta t}{2}\right) = v\left(t - \frac{\Delta t}{2}\right) + a_{av}\Delta t \quad (4.6)$$

Εάν πάλι θεωρηθεί ότι η επιτάχυνση είναι πολύ κοντά σε γραμμική μορφή για το διάστημα χρόνου από $\left(t - \frac{\Delta t}{2}\right)$ έως $\left(t + \frac{\Delta t}{2}\right)$ τότε η μέση τιμή της επιτάχυνσης είναι η στιγμιαία τιμή της στον χρόνο t .



$$a_{av} = a(t) \quad (4.7)$$

και από την οποία εξάγεται η ακόλουθη εξίσωση για την ταχύτητα.

$$v\left(t + \frac{\Delta t}{2}\right) = v\left(t - \frac{\Delta t}{2}\right) + a(t)\Delta t \quad (4.8)$$

Συνδυάζοντας τις εξισώσεις (4.4) και (4.8) προκύπτει μία απλή εξίσωση για την ενημέρωση των συντεταγμένων.

$$R\left(t + \frac{\Delta t}{2}\right) = R(t) + v\left(t + \frac{\Delta t}{2}\right)\Delta t \quad (4.9)$$

Οι εξισώσεις (4.8) και (4.9) καλούνται ως leapfrog μέθοδοι επειδή η ταχύτητα είναι εκτός βήματος, ενώ οι αποστάσεις είναι στο βήμα του μισού χρόνου.

Αυτό είναι και το σημαντικότερο μειονέκτημα της μεθόδου, το ότι δηλαδή οι θέσεις και οι ταχύτητες που υπολογίζονται είναι κατά μισό βήμα χρόνου, εκτός συγχρονισμού. Συνοπτικά ο κύκλος που ακολουθείται για να υπολογισθεί κάποια επόμενη συντεταγμένη σε κάποια χρονική στιγμή σύμφωνα με την παραπάνω μέθοδο είναι:

1. Υπολογίζεται η επιτάχυνση $-\frac{1}{m} \frac{dV}{dR}$ στο χρόνο t .
2. Ενημερώνεται η ταχύτητα στο χρόνο $\left(t + \frac{\Delta t}{2}\right)$ λαμβάνοντας υπόψη την τιμή που είχε στον χρόνο $\left(t - \frac{\Delta t}{2}\right)$, με χρήση της εξίσωσης (4.8).
3. Ενημερώνονται οι συντεταγμένες για το χρόνο $(t + \Delta t)$ σχετικά με την τιμή που είχαν στον χρόνο t σύμφωνα με την εξίσωση (4.9).

4.3.2.2. Ολοκληρωτής ταχυτήτων του Verlet

Ο αλγόριθμος αυτός είναι παρόμοιος με τον προηγούμενο και συνοπτικά λειτουργεί όπως παρακάτω:

Με δεδομένη θέση $R(t)$, ταχύτητα $v(t)$ και επιτάχυνση $a(t)$ σε δεδομένο χρόνο t μπορούν να υπολογισθούν οι ταχύτητα, θέση και επιτάχυνση σε μεταγενέστερη χρονική στιγμή ως εξής:

$$a(t + \Delta t) = \frac{F(t + \Delta t)}{m} \quad (4.10)$$

$$R(t + \Delta t) = R(t) + \Delta t v(t) + \frac{\Delta t^2 a(t)}{2} \quad (4.11)$$



$$v(t + \Delta t) = v(t) + \frac{1}{2} \Delta t [a(t) + a(t + \Delta t)] \quad (4.12)$$

Οι αρχικές καρτεσιανές συντεταγμένες των ατόμων του μοριακού συστήματος είναι καθορισμένες σε κάποιο αρχείο εισόδου.

Οι αρχικές ταχύτητες έχουν δημιουργηθεί τυχαία στην αρχή μιας δυναμικής πορείας σύμφωνα με τη θερμοκρασία και ένα τυχαίο γεννήτορα αριθμό.

Για να επαναληφθεί η ίδια ακριβώς δυναμική πορεία είναι απαραίτητο να ορισθεί ο ίδιος τυχαίος γεννήτορας αριθμός που είχε χρησιμοποιηθεί στην προηγούμενη πορεία.

4.3.2.3. Επιλογή του βήματος χρόνου.

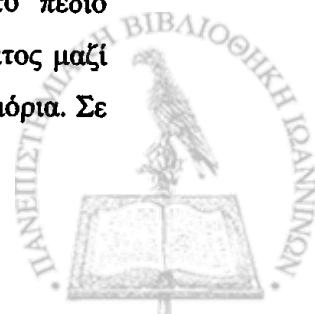
Μία σημαντική παράμετρος κατά την ολοκλήρωση του αλγορίθμου είναι η επιλογή του βήματος Δt . Για την καλύτερη χρήση του χρόνου του υπολογιστή θα πρέπει να χρησιμοποιηθεί μεγάλο βήμα χρόνου ολοκλήρωσης. Ωστόσο όσο μεγαλύτερο είναι το βήμα του χρόνου τόσο μεγαλύτερη αστάθεια και ανακρίβεια προκαλείται στη διαδικασία της ολοκλήρωσης. Το βήμα του χρόνου που χρησιμοποιείται κάθε φορά, εξαρτάται από το σύστημα καθώς και από τους ολοκληρωτές που θα χρησιμοποιηθούν. Ένας βασικός περιορισμός που επιβάλλεται από το μοριακό σύστημα είναι η μέγιστη τιμή της συχνότητας κίνησης των ατόμων, η οποία πρέπει να ληφθεί υπόψη.

Μία παλμική περίοδος των μοριακών συστημάτων για να ικανοποιεί τη υπόθεση του Verlet πρέπει να χωριστεί σε 8-10 τμήματα, έτσι ώστε οι ταχύτητες και οι επιταχύνσεις να είναι συνεχείς στο βήμα του χρόνου που χρησιμοποιείται.

Στα περισσότερα μοριακά συστήματα η υψηλότερη παλμική συχνότητα είναι αυτή του δεσμού C-H του οποίου η περίοδος είναι της τάξης του 10^{-14} sec. Το βήμα του χρόνου ολοκλήρωσης σε αυτή την περίπτωση θα πρέπει να είναι περίπου $1 \text{ fs} = 10^{-15}$ sec. Ανάλογα με τη μέθοδο ολοκλήρωσης που χρησιμοποιείται πρέπει να επιλέγεται και το κατάλληλο βήμα χρόνου ολοκλήρωσης. Το βήμα του χρόνου που συνίσταται για τη leapfrog μέθοδο και τη μέθοδο των ταχυτήτων του Verlet κυμαίνεται από 0.1 – 1 fs.

4.3.2.4. Περιοδικές Οριακές Συνθήκες

Η συνεκτικές δυνάμεις μεταξύ των μορίων που περιγράφονται από το πεδίο δυνάμεων, μπορεί να είναι ικανές ώστε να κρατούν τα μόρια του συστήματος μαζί κατά τη διάρκεια μιας προσομοίωσης. Αυτό ισχύει σε ένα σύστημα με λίγα μόρια. Σε



μεγαλύτερα συστήματα N μορίων εφαρμόζεται η μέθοδος των περιοδικών οριακών συνθηκών (Periodic Boundary Conditions).

Σύμφωνα με τη μέθοδο αυτή ένας κύβος αντιγράφεται στον χώρο δημιουργώντας ένα απέραντο πλέγμα. Κατά τη διάρκεια της μοριακής δυναμικής προσομοίωσης, κάθε μόριο κινείται μέσα στον χώρο ενός τέτοιου κύβου και ένα περιοδικό αντίγραφο του κινείται με το ίδιο τρόπο στον χώρο κάθε ενός από τους γειτονικούς κύβους. Έτσι καθώς ένα μόριο, εξαιτίας της κίνησης του, φύγει από τον αρχικό κύβο τότε κάποιο άλλο αντίγραφο του εισέρχεται σε αυτόν από την αντίθετη πλευρά. Ο κύβος απλά αναπαριστά ένα βολικό σύστημα αξόνων για τις μετρήσεις των συντεταγμένων των N μορίων.

4.3.3. Θερμοκρασία

Η έννοια της θερμοκρασίας είναι θεμελιώδεις προκειμένου για προσομοιώσεις μοριακής δυναμικής. Γενικά η θερμοκρασία είναι ανάλογη της κινητικής ενέργειας ενός συστήματος, και ισοδύναμη με τις ατομικές ταχύτητες του συστήματος. Η σχέση μεταξύ θερμοκρασίας και ταχύτητας των ατόμων αποδεικνύεται από την κινητική θεωρία των αερίων. Η θεμελιακή σχέση θερμοκρασίας και ταχυτήτων των ατόμων είναι η εξίσωση Maxwell - Boltzman (MB).

$$f(v)dv = \left(\frac{m}{2\pi kT} \right)^{\frac{3}{2}} e^{-\frac{mv^2}{2kT}} 4\pi v^2 dv \quad (4.13)$$

Αυτή η πολύ γνωστή εξίσωση εκφράζει την πιθανότητα $f(v)$ για ένα άτομο μάζας m να έχει ταχύτητα v όταν η θερμοκρασία είναι T . Στην πραγματικότητα μια απλή ταχύτητα δεν μπορεί να εκφραστεί με μία θερμοκρασία. Η θερμοκρασία είναι το μέτρο μιας μέσης τιμής της κατανομής και έχει νόημα σε αυστηρά θερμοδυναμική αίσθηση και μόνο σε ισορροπία του συστήματος. Στην πραγματικότητα η εξίσωση MB καθορίζει πως πρέπει να είναι αυτή η κατανομή. Παρ' όλα αυτά η εξίσωση MB μπορεί να χρησιμοποιηθεί για να δώσει λύσεις σε πολλές μέσες τιμές ταχυτήτων. Για παράδειγμα η πιο πιθανή ταχύτητα v_p είναι:

$$v_p = \left(\frac{2kT}{m} \right)^{\frac{1}{2}} \quad (4.14)$$



Η ταχύτητα είναι ανάλογη της τετραγωνικής ρίζας της θερμοκρασίας, όπως αναμένεται, εφ' όσον η θερμοκρασία είναι ανάλογη της κινητικής ενέργειας η οποία είναι ίση με $\frac{1}{2}mv^2$.

4.3.3.1. Πως υπολογίζεται η θερμοκρασία.

Η θερμοκρασία είναι μία θερμοδυναμική ποσότητα η οποία είναι έχει νόημα μόνο στην κατάσταση ισορροπίας και μπορεί να συσχετισθεί με τη μέση κινητική ενέργεια του συστήματος ως εξής:

Κάθε βαθμός ελευθερίας (είτε της ορμής είτε των συντεταγμένων) που εμφανίζεται σαν τετραγωνικός όρος στην Χαμιλτωνιανή έχει μία μέση ενέργεια ίση με $KT/2$ συνδεδεμένη μαζί του. Αυτό ισχύει για τις ορμές p_i οι οποίες εμφανίζονται στην Χαμιλτωνιανή ως $p_i^2/2m$.

$$\left\langle \sum_i^N \frac{p_i^2}{2m} \right\rangle = \langle K \rangle = \frac{N_f K_B T}{2} \quad (4.15)$$

Η αριστερή πλευρά της εξίσωσης (4.15) ονομάζεται μέση κινητική ενέργεια του συστήματος, N_f είναι ο αριθμός των βαθμών ελευθερίας και T είναι η θερμοδυναμική θερμοκρασία. Σε ένα απεριόριστο σύστημα με N άτομα το N_f είναι ίσο με $3N$ επειδή το κάθε άτομο έχει τρεις συνιστώσες ταχύτητας, τις v_x, v_y, v_z .

Η στιγμιαία θερμοκρασία ορίζεται ως εξής:

$$T_{ins} = \frac{2K}{N_f k_B} \quad (4.16)$$

Ο μέσος όρος της στιγμιαίας θερμοκρασίας T_{ins} είναι η θερμοδυναμική θερμοκρασία. Η θερμοκρασία υπολογίζεται από την συνολική κινητική ενέργεια και από το συνολικό αριθμό των βαθμών ελευθερίας. Για ένα μη περιοδικό σύστημα ισχύει:

$$\frac{(3N-6)k_B T}{2} = \sum_{i=1}^N \frac{m_i v_i^2}{2} \quad (4.17)$$

Έχουν αφαιρεθεί έξι βαθμοί ελευθερίας επειδή αγνοήθηκε η μεταφορά και η περιστροφή του κέντρου μάζας. Σε ένα περιοδικό σύστημα ισχύει:

$$\frac{(3N-3)k_B T}{2} = \sum_{i=1}^N \frac{m_i v_i^2}{2} \quad (4.18)$$

Μόνο οι τρεις βαθμοί ελευθερίας που ανταποκρίνονται στη μεταφορική κίνηση μπορούν να αγνοηθούν αφού η περιστροφή ενός κεντρικού ατόμου επιβάλλει ροπή στρέψης στα γειτονικά άτομα.



4.3.4. Στάδια και διάρκεια των προσομοιώσεων της μοριακής δυναμικής.

Οι δυναμικές προσομοιώσεις συνήθως διεξάγονται σε δύο στάδια

1. Το στάδιο εξισορρόπησης του συστήματος.
2. Το στάδιο ισορροπίας και συλλογής δεδομένων¹⁸².

Η διάρκεια του κάθε σταδίου εξαρτάται από το μοριακό σύστημα όπως επίσης και από τις ιδιότητες του συστήματος που έχουν βρεθεί μετά από έρευνα. Στη συνέχεια περιγράφεται το στάδιο 1, η καλή εκτέλεση του οποίου έχει επιπτώσεις στην ποιότητα των δεδομένων του σταδίου 2.

4.3.4.1. Το στάδιο εξισορρόπησης

Ο σκοπός της εξισορρόπησης είναι να προετοιμάσει το σύστημα έτσι ώστε αυτό να καταλήξει στην πιο πιθανή διαμόρφωση, σύμφωνα με τη στοχευόμενη θερμοκρασία και πίεση. Όταν ένα μοριακό σύστημα είναι μεγάλης κλίμακας, συνήθως δαπανά μεγάλο υπολογιστικό χρόνο μέχρι να φτάσει στην κατάσταση ισορροπίας, εξαιτίας του τεράστιου διαμορφωσιακού χώρου που υπάρχει για να ερευνηθεί.

Οι ισοδυναμικές επιφάνειες της ενέργειας είναι ένας άλλος παράγοντας που πρέπει να ληφθεί υπόψη. Εάν τα ενεργειακά φράγματα ανάμεσα σε τοπικά ελάχιστα και στο γενικό ελάχιστο, είναι υψηλά τότε το πέρασμα του κάθε φράγματος είναι δύσκολο ή και αδύνατο με αποτέλεσμα να σπαταληθεί περισσότερος υπολογιστικός χρόνος.

Ένας τρόπος που κρίνει το εάν ένα σύστημα έχει ισορροπήσει είναι η γραφική αναπαράσταση των τις διαφόρων θερμοδυναμικών ποσοτήτων όπως είναι η ενέργεια, η θερμοκρασία και η πίεση σε σχέση με τον χρόνο και η παρατήρηση της σταθεροποίησης των ποσοτήτων αυτών. Όταν επέλθει ισορροπία αυτές οι ποσότητες κυμαίνονται γύρω από το μέσο όρο τους ο οποίος παραμένει συνεχής και σταθερός συναρτήσει του χρόνου. Αυτός είναι απαραίτητος αλλά όχι ικανοποιητικός έλεγχος, επειδή δεν είναι ασυνήθιστο να συμβεί μία ξαφνική διαμορφωσιακή αλλαγή μετά από ένα μεγάλο χρονικό διάστημα. Ένας άλλος τρόπος ελέγχου της ισορροπίας είναι η εκκίνηση του υπολογισμού με διαφορετικές αρχικές διαμορφώσεις και διαφορετικές αρχικές ταχύτητες.

Η σύγκλιση σε παρόμοιες διαμορφώσεις και ιδιότητες, ενώ η έναρξη έγινε από διαφορετικές αρχικές συνθήκες, είναι ένας καλός δείκτης του ότι έχει επιτευχθεί η ισορροπία στο σύστημα.



Έχει υιοθετηθεί κατά τη διάρκεια των πρώτων βημάτων της δυναμικής προσομοίωσης, όταν το σύστημα είναι μακριά από την ισορροπία, οι ταχύτητες να κλιμακώνονται ισχυρά έτσι ώστε να επιταχύνεται η σύγκλιση προς στην επιθυμητή θερμοκρασία.

Κάθε φορά που η τρέχουσα τιμή της θερμοκρασίας είναι υψηλότερη ή χαμηλότερη από το στόχο, καθορίζεται μία τιμή κλιμάκωσης, που συνήθως παίρνει την τιμή των 10 K. Αυτή προσθέτει ή αφαιρεί ενέργεια από το σύστημα αποτελεσματικά.

Το βασικό όριο για να επιτευχθεί ισορροπία είναι το πόσο γρήγορα μπορεί να μεταφερθεί η ενέργεια από, προς και ανάμεσα στους διάφορους εσωτερικούς βαθμούς ελευθερίας του μορίου.

Η ταχύτητα αυτής της επεξεργασίας εξαρτάται από την έκφραση του Δυναμικού της Ενέργειας, τις παραμέτρους και τη φύση των παλμικών, περιστροφικών και παράλληλων μετατοπίσεων του συστήματος.

Επίσης εξαρτάται άμεσα από το μέγεθος του συστήματος. Τα μεγαλύτερα συστήματα απαιτούν περισσότερο χρόνο μέχρι να ισορροπήσουν.



4.4. Ab-initio μοριακή δυναμική

Οι ab-initio μέθοδοι μοριακής δυναμικής αποτελούν μία εξαιρετική επέκταση των κλασικών μεθόδων που στηρίζονται σε διάφορα παραμετροποιημένα πεδία δυνάμεων (force fields). Η κεντρική ιδέα σε κάθε ab-initio μοριακή δυναμική στηρίζεται στον “on-the-fly” υπολογισμό των δυνάμεων (πρώτων παραγώγων), που δρουν στους πυρήνες των ατόμων, καθώς το μοριακό σύστημα εξελίσσεται στο χρόνο και δημιουργείται η τροχιά. Με αυτήν την θεώρηση το πρόβλημα μεταφέρεται στην επίλυση της κβαντικής εξίσωσης Schrodinger.

Οι εφαρμογές της ab-initio μοριακής δυναμικής είναι τόσες πολλές, γεγονός που επιβεβαιώνεται από την δημιουργία ξεχωριστού αριθμού ταξινόμησης PACS το 1996 με τον κωδικό 71.15.Pd «Electronic Structure: Molecular dynamics calculations (Car-Parrinello) and other numerical simulations» που εισήχθη στο Physics and Astronomy Classification Scheme¹⁸³.

Το πλεονέκτημα των ab-initio προσομοιώσεων είναι ότι παράγουν μεγάλη ποσότητα πληροφορίας η οποία δεν είναι διαθέσιμη από τις παραμετροποιημένες μεθόδους με τα πεδία δυνάμεων (κλασική μηχανική). Ένα άλλο πολύ σημαντικό χαρακτηριστικό των προσομοιώσεων μοριακής δυναμικής με ab-initio μεθόδους είναι η δυνατότητα να μελετηθούν μοριακά συστήματα που εμπλέκουν διάσπαση ή σχηματισμό δεσμών μεταξύ των ατόμων. Η διάσπαση και ο σχηματισμό δεσμών απαιτεί υπολογιστικές μεθόδους ικανές να αντιμετωπίσουν πολλές ηλεκτρονικές καταστάσεις. Η κλασική μοριακή δυναμική υστερεί, ενώ στις περισσότερες περιπτώσεις αδυνατεί να περιγράψει αυτά τα φυσικά φαινόμενα.

Αυτά όμως τα πλεονεκτήματα ισοσταθμίζονται από το πολύ μεγάλο υπολογιστικό κόστος των ab-initio προσομοιώσεων μοριακής δυναμικής. Στην περίπτωση που ένα μοριακό σύστημα αποτελείται από N άτομα το υπολογιστικό κόστος είναι της τάξης του 10^N .

Οι πιο διαδεδομένες μέθοδοι ab-initio προσομοιώσεων μοριακής δυναμικής είναι η μέθοδος Born-Oppenheimer (BO) και η μέθοδος Car-Parrinello¹⁸⁴ (CP). Η εφαρμογή και ανάπτυξη μιας υπολογιστικής διαδικασίας BO είναι το αντικείμενο του παρόντος κεφαλαίου.

Η μέθοδος BO έχει το πλεονέκτημα ότι δίνει μία πολύ καλή περιγραφή, μιας καλά ορισμένης επιφάνειας δυναμικής ενέργειας, ενώ η μέθοδος CP προσφέρει σχεδόν την ίδια ακρίβεια αλλά με λιγότερο υπολογιστικό κόστος¹⁸⁵. Ένα άλλο σημαντικό πλεονέκτημα της μεθόδου BO είναι ότι μπορεί να εφαρμοστεί σε μεγάλη ποικιλία



κυματοσυναρτήσεων και μεθόδων όπως MCSCF, DFT, MP2 κ.α., παρέχοντας την πιο επακριβή περιγραφή της υπερεπιφάνειας δυναμικής ενέργειας^{186,187} ενός μοριακού συστήματος.

4.4.1. Born-Oppenheimer μοριακή δυναμική

Η πιο κλασική μέθοδος ab-initio μοριακής δυναμικής, βασίζεται στην προσέγγιση BO. Στην μέθοδο αυτή, η ηλεκτρονική δομή υπολογίζεται σε κάθε σημείο της τροχιάς με την επίλυση της χρονικά ανεξάρτητης εξίσωσης Schrodinger (π.χ. μέθοδος Hartree Fock), ενώ οι τροχιά των ατόμων αντιμετωπίζεται με την κλασική προσέγγιση.

Οι εξισώσεις που περιγράφουν αυτή την μέθοδο είναι:

$$m_i \frac{\partial^2 R_i}{\partial t_i^2} = -\nabla_i \min \langle \Psi_0 | H_e | \Psi_0 \rangle \quad (4.19)$$

και:

$$E_0 \Psi_0 = H_e \Psi_0 \quad (4.20)$$

Στην πράξη οι εξισώσεις (4.19) και (4.20) άρχισαν να εφαρμόζονται ευρέως σε προσομοιώσεις μοριακής δυναμικής από τις αρχές της δεκαετίας του 1990¹⁸⁸, στηριζόμενες στην εξέλιξη των πακέτων κβαντικών υπολογισμών με την θεωρία Hartree-Fock σε ab-initio και semi-empirical υπολογισμούς.

Σε κάθε βήμα της μοριακής δυναμικής οι εξισώσεις (4.19) και (4.20) επιλύονται και το υπολογιστικό κόστος ανάγεται στην επίλυση των εξισώσεων με την μέθοδο Hartree-Fock.

Η δυναμική εξέλιξη του μοριακού συστήματος εξαρτάται:

- a. Από τον υπολογισμό των δυνάμεων (πρώτων παραγώγων) από την εξίσωση (4.19) ως προς τις θέσεις των ατόμων. Το πλεονέκτημα είναι ότι σχεδόν όλες οι κυματοσυναρτήσεις που στηρίζονται στην θεωρία Hartree-Fock παρέχουν αναλυτικές παραγώγους της ενέργειας ως προς τις ατομικές συντεταγμένες.
- b. Από τον υπολογισμό του βήματος ολοκλήρωσης από την εξίσωση (4.10). Επειδή τα βήματα ολοκλήρωσης για την επίλυση των εξισώσεων κίνησης των ατόμων είναι πάρα πολλά, έχουμε μεγάλο υπολογιστικό κόστος.



4.5. Η διασύνδεση TINKER και GAMESS-US

Σε αυτό το κεφάλαιο περιγράφεται η ανάπτυξη ενός υπολογιστικού συστήματος ικανού να εκτελεί *ab-initio* προσομοιώσεις μοριακής δυναμικής, βασισμένη στην προσέγγιση BO. Η κεντρική ιδέα για την ανάπτυξη ενός γενικής χρήσης συστήματος είναι η εξής: *“Το GAMESS-US παρέχει τις δυνάμεις (forces) και η ολοκλήρωση των εξισώσεων κίνησης του Newton στον χρόνο γίνεται από το TINKER”*. Μια παρόμοια διαδικασία αναπτύχθηκε¹⁸⁹ το 1996 με την διασύνδεση των πακέτων VENUS και MOPAC.

Για την υλοποίηση αυτού του συστήματος χρησιμοποιήθηκε η τεχνολογία που περιγράφηκε στο κεφάλαιο 2 ενώ κώδικας αναπτύχθηκε για το πακέτο TINKER. Αρχικά αναπτύχθηκε η βιβλιοθήκη διασύνδεσης *“gms_api.c”* του TINKER με το GAMESS-US. Ο κώδικας της βιβλιοθήκης συνδέθηκε με την εφαρμογή μοριακής δυναμικής *“dynamicab”* που δημιουργήθηκε χρησιμοποιώντας τις ρουτίνες ανάπτυξης του TINKER.

Το τελικό σύστημα διαθέτει όλα τα χαρακτηριστικά μοριακής δυναμικής που παρέχει το TINKER και εκτελείται αυτόνομα με εκτέλεση της εφαρμογής *“dynamicab”*, αφού πρώτα δημιουργηθούν όλα τα κατάλληλα αρχεία εισόδου και οριστούν οι επιμέρους παράμετροι. Ο χρήστης δημιουργεί ένα αρχείο εισόδου με τις καρτεσιανές συντεταγμένες του μοριακού συστήματος, προδιαγεγραμμένο για υπολογισμό με το σύστημα GAMESS-US και η μοριακή δυναμική εκτελείται από την εφαρμογή *“dynamicab”*. Στην συνέχεια θα περιγράψουμε την αρχιτεκτονική, τις λειτουργικές προδιαγραφές και τις αναλυτικές τεχνικές λεπτομέρειες της ανάπτυξης του συστήματος.

4.5.1. Η αρχιτεκτονική του συστήματος

Η ανάπτυξη ενός σταθερού και διαλειτουργικού συστήματος επικοινωνίας μεταξύ των πακέτων GAMESS-US και TINKER έγινε με την ανάπτυξη κώδικα σε γλώσσα προγραμματισμού ANSI-C. Ο κώδικας αναπτύχθηκε ως βιβλιοθήκη διασύνδεσης (library) μεταξύ των δύο πακέτων και συνδέεται με τον εκτελέσιμο κώδικα της εφαρμογής *“dynamicab”* μέσω των ρουτίνων ανάπτυξης εφαρμογών που παρέχει το TINKER.

Ο πηγαίος κώδικας της βιβλιοθήκης που αναπτύχθηκε και είναι ο βασικό μηχανισμός λειτουργίας του συστήματος, βρίσκεται στο αρχείο *“gms_api.c”* που διανέμεται με το σύστημα. Οι συναρτήσεις που αναπτύχθηκαν έχουν ως κύριο στόχο:



- a. Την διαχείριση του αρχείου εισόδου (input) του πακέτου GAMESS-US.
- b. Την απρόσκοπτη δι-επικοινωνία της εφαρμογής "dynamicab" με το GAMESS-US μέσω κλήσεων σε διαδικασίες του λειτουργικού συστήματος.

Η αναλυτική αρχιτεκτονική του συστήματος καθώς και η αλληλουχία κλήσεων φαίνεται στην εικόνα 4.1.

Στην αρχιτεκτονική του συστήματος (εικόνα 4.1) φαίνεται το κύριο σύστημα διαχείρισης που είναι η εφαρμογή "dynamicab". Η επικοινωνία και ανταλλαγή δεδομένων μεταξύ του TINKER και GAMESS-US γίνεται με την χρήση C-wrappers, όπως περιγράφηκαν στο κεφάλαιο 2 (ο κώδικας που αναπτύχθηκε σε C χωρίστηκε σε δύο λογικά υποσύνολα συναρτήσεων C-wrappers και C-primitives).

Για το λόγο αυτό το TINKER και πιο συγκεκριμένα η εφαρμογή "dynamicab" επικοινωνεί μόνο μέσω των C-wrappers με το GAMESS-US παρέχοντας έμπιστο και ασφαλή τρόπο διεπικοινωνίας και ανταλλαγής δεδομένων.

Όλες οι συναρτήσεις C-wrappers παρέχουν ασφαλή κώδικα διεπαφής μεταξύ των δυο πακέτων και βρίσκονται στο πηγαίο αρχείο "dyn_ext.h", υποστηρίζοντας όλους τους γνωστούς μεταγλωττιστές της γλώσσας προγραμματισμού C. Το υποσύνολο συναρτήσεων C-primitives, καλείται μόνο από συναρτήσεις C-wrappers. Αφορά την εκτέλεση λειτουργιών που σχετίζονται με την εκτέλεση του GAMESS-US και την πρόσβαση στις εξωτερικές δομές αποθήκευσης πληροφοριών που αφορούν τους κβαντικούς μοριακούς υπολογισμούς. Η πιο σημαντική συνάρτηση του συνόλου των C-primitives είναι η συνάρτηση GameSSProperty στην οποία καταλήγει η αρχιτεκτονική του συστήματος και παρέχει τα αποτελέσματα των κβαντικών υπολογισμών στους C-wrappers ώστε με την σειρά τους να «περάσουν» στην εφαρμογή "dynamicab". Η αναλυτική περιγραφή των ρουτινών C-primitives και C-wrappers βρίσκεται στα παραρτήματα Α και Β αντίστοιχα.

Για να επιτευχθεί η σύνδεση και η δι-επικοινωνία, πρέπει το GAMESS-US να γνωρίζει ποια αρχεία θα χρησιμοποιήσει και η εφαρμογή "dynamicab" ποια αρχεία αποτελεσμάτων θα διαβάσει. Σε αυτό το σημείο γίνεται χρήση των μεταβλητών περιβάλλοντος (environment variables) που διαθέτει κάθε λειτουργικό σύστημα. Για το λόγο αυτό ένα περιγραφικό αρχείο κελύφους (shell script) χρησιμοποιείται για την εκκίνηση των προσομοιώσεων μοριακής δυναμικής.

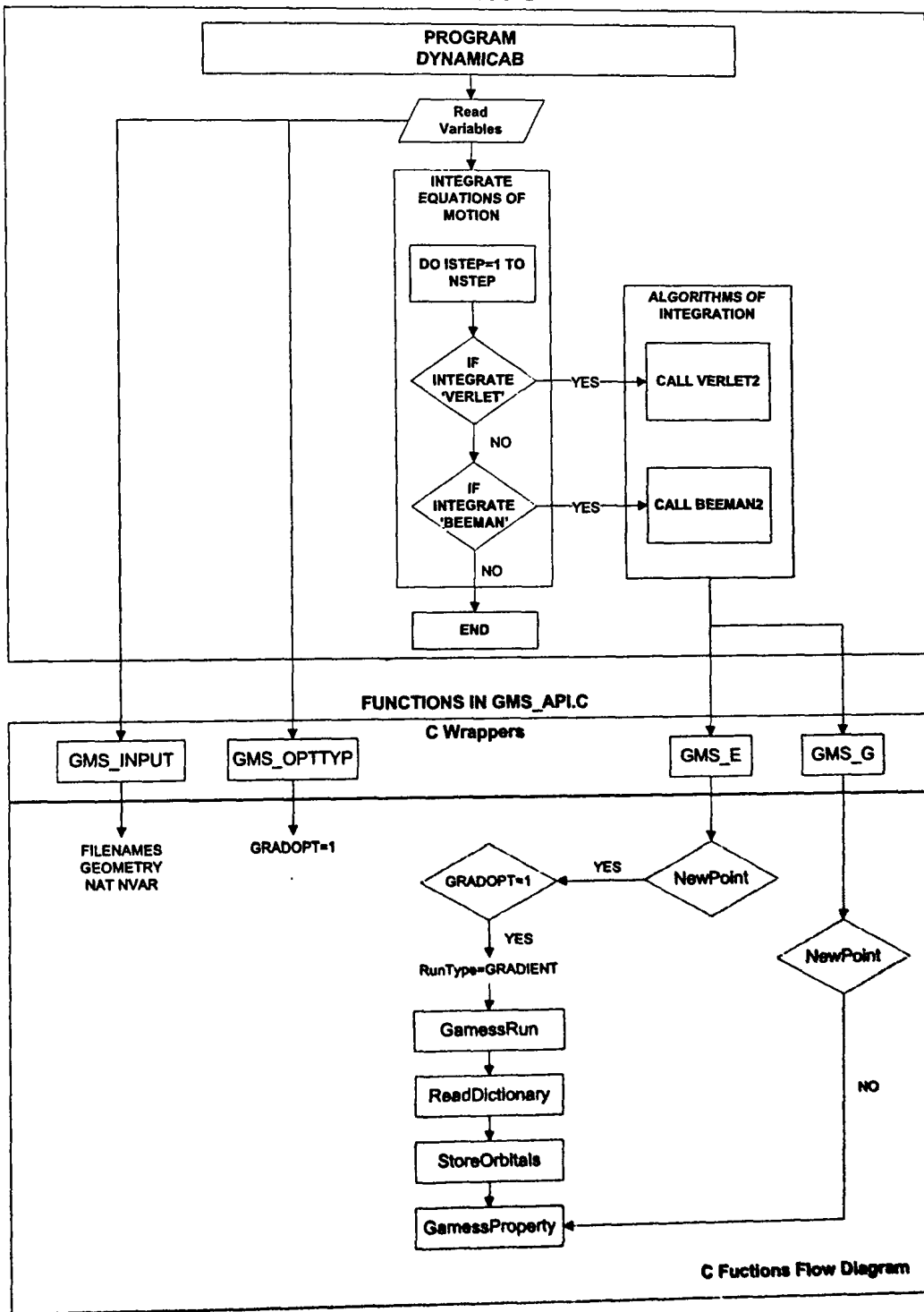
Το αρχείο αυτό περιέχει σε μεταβλητές περιβάλλοντος τα ονόματα και την θέση των αρχείων που θα χρησιμοποιηθούν για την επικοινωνία. Κάθε πρόγραμμα που εκτελείται σε ένα λειτουργικό σύστημα έχει πρόσβαση στις μεταβλητές



περιβάλλοντος τις οποίες διαβάζει ώστε να πάρει πληροφορίες για εξωτερικά δεδομένα που σχετίζονται με την εκτέλεσή του.

Εικόνα 4.1

TINKER TOOLS



Το σύστημα που αναπτύχθηκε έχει την δυνατότητα να εκτελέσει προσομοιώσεις μοριακής δυναμικής με όλες τις κυματοσυναρτήσεις που παρέχει το GAMESS-US, γεγονός που το καθιστά ένα πολύ ισχυρό εργαλείο για την μελέτη μεγάλων μοριακών συστημάτων. Έχει την δυνατότητα να λειτουργεί με όλες τις τρέχουσες εκδόσεις του GAMESS-US όπως και με παραλλαγές του.

4.5.2. Οι ρουτίνες που αναπτύχθηκαν στο πακέτο TINKER

Χρησιμοποιώντας την βιβλιοθήκη ανάπτυξης εργαλείων του TINKER αναπτύχθηκαν οι εξής ρουτίνες:

- a. Η κυρίως εφαρμογή *ab-initio* μοριακής δυναμικής με το όνομα “dynamicab”.
- b. Η ρουτίνα BEEMAN2 για την υλοποίηση ενός βήματος μοριακής δυναμικής με την μέθοδο ολοκλήρωσης “Better Beeman”.
- c. Η ρουτίνα VERLET2 για την υλοποίηση ενός βήματος μοριακής δυναμικής με την μέθοδο ολοκλήρωσης “velocity Verlet”.

4.5.3. Η εφαρμογή dynamicab

Η εφαρμογή *dynamicab* αναπτύχθηκε σε Fortran 77 κατ’ αναλογία με την εφαρμογή “dynamic” που παρέχει το TINKER για προσομοιώσεις μοριακής δυναμικής. Σκοπός είναι η ολοκλήρωση των εξισώσεων κίνησης, καλώντας τις υπορουτίνες BEEMAN2 ή VERLET2, ανάλογα με την επιλογή του χρήστη. Η διαδικασία γίνεται επαναληπτικά για όλο τον χρόνο της προσομοίωσης. Οι αλγόριθμοι BEEMAN2 και VERLET2 βρίσκονται σε ξεχωριστά αρχεία και κάνουν κλήσεις στο πακέτο GAMESS-US για τον υπολογισμό: α) της δυναμικής ενέργειας (*ab-initio* ηλεκτρονική ενέργεια) καθώς και β) των πρώτων παραγώγων που χρησιμοποιούνται για τον υπολογισμό των δυνάμεων.

Οι C-wrappers συναρτήσεις GMS_OPTTYP και GMS_INPUT χρησιμοποιούνται αρχικά για να αποκατασταθεί η επικοινωνία με το GAMESS-US ως εξής:

- Η συνάρτηση GMS_OPTTYP ορίζει την δυνατότητα να υπολογίζονται αναλυτικά οι πρώτες παράγωγοι της ηλεκτρονικής (δυναμικής) ενέργειας του συστήματος σε κάθε βήμα της μοριακής δυναμικής από το GAMESS-US. Δηλαδή κάθε υπολογισμός να είναι τύπου GRADIENT.
- Η συνάρτηση GMS_INPUT ορίζει τις μεταβλητές περιβάλλοντος και τα κοινά αρχεία επικοινωνίας των δύο εφαρμογών.



Εκτός από τις C-wrappers συναρτήσεις η εφαρμογή “dynamicab” κάνει χρήση των επόμενων ρουτινών που παρέχει η βιβλιοθήκη ανάπτυξης εφαρμογών του TINKER:

**BEEMAN2 FINAL GETXYZ INITIAL MDINIT MDREST MECHANIC NEXTARG
RGDSTEP SDSTEP SHAKEUP VERLET2**

Οι ρουτίνες BEEMAN και VERLET αντικαταστάθηκαν από τις νέες ρουτίνες BEEMAN2 και VERLET2 αντίστοιχα, που αναπτύχθηκαν για τις ab-initio προσομοιώσεις μέσω του GAMESS-US όπως φαίνεται και στην εικόνα 4.1. Ο πηγαίος κώδικας της εφαρμογής “dynamicab” βρίσκεται στο αρχείο “dynamicab.f”.

4.5.4. Οι ρουτίνες BEEMAN2 και VERLET2

Οι ρουτίνες BEEMAN2 και VERLET2 αναπτύχθηκαν σε Fortran 77 χρησιμοποιώντας τις συναρτήσεις C-wrappers GMS_E και GMS_G της βιβλιοθήκης διασύνδεσης καθώς και τις συναρτήσεις του TINKER:

KINETIC MDSAVE MDSTAT PRESSURE RATTLE RATTLE2 TEMPER TEMPER2

Σκοπός των δύο αυτών συναρτήσεων είναι η επιλογή του επόμενου βήματος κατά την χρονική εξέλιξη της προσομοίωσης της ab-initio μοριακής δυναμικής.

Οι συναρτήσεις αυτές είναι το ακριβές ανάλογο της κλασσικής μοριακής δυναμικής με την διαφορά ότι οι δυνάμεις υπολογίζονται από το πακέτο GAMESS-US μέσω των συναρτήσεων GMS_E και GMS_G. Είναι φανερό ότι σε κάθε βήμα (εικόνα 4.1) γίνεται κλήση στο GAMESS-US το οποίο εκτελεί τον κβαντικό υπολογισμό για το τρέχον σημείο της τροχιάς. Συνολικά γίνονται NSTEP κλήσεις στις ρουτίνες BEEMAN2 ή VERLET2, που σημαίνει ότι το GAMESS-US εκτελείται αντίστοιχα NSTEP φορές.

Η συνάρτηση GMS_E επιστρέφει κατά διεύθυνση την ολική ηλεκτρονική ενέργεια (δυναμική ενέργεια) και η GMS_G τις πρώτες αναλυτικές παραγώγους από τις οποίες υπολογίζεται η δύναμη (force) που ασκείται στα άτομα του μοριακού συστήματος.

Στην συνέχεια δίνουμε ένα τμήμα του κώδικα της VERLET2 σε Fortran 77 όπου γίνεται ο υπολογισμός ενός βήματος της μοριακής δυναμικής:

```
C      get potential energy
C      call GMS_E(XX,En)
C      get potential energy
C      call GMS_E(XX,En)
C      get analytic derivatives of potential energy
C      call GMS_G(XX,derive)
```



```
c   use Newton's second law to get the next accelerations
c   find the full-step velocities using the Verlet recursion
  do i = 1, n
    if (use(i)) then
      do j = 1, 3
        a(j,i) = -convert * derivs(j,i) / mass(i)
        v(j,i) = v(j,i) + a(j,i)*dt_2
      end do
    end if
  end do
```

Ο πηγαίος κώδικας των υπορουτινών BEEMAN2 και VERLET2 με σχόλια βρίσκεται στα αρχεία "beeman2.f" και "verlet2.f" και παρουσιάζεται στα παραρτήματα Η και Θ αντίστοιχα. Οι ρουτίνες αυτές διασυνδέονται (linking) με την εφαρμογή "dynamicab" για την παραγωγή του εκτελέσιμου κώδικα της ab-initio μοριακής δυναμικής. Η κλήση στο GAMESS-US όπως φαίνεται και στο σχήμα της αρχιτεκτονικής (εικόνα 4.1) γίνεται μόνο μέσω αυτών των 2 ρουτινών.

Στην επόμενη ενότητα των εφαρμογών θα φανεί πως το σύστημα που αναπτύχθηκε εφαρμόζεται σε προβλήματα βιολογικής σημασίας καθώς και οι τεχνικές οδηγίες για την εκτέλεση μιας προσομοίωσης.



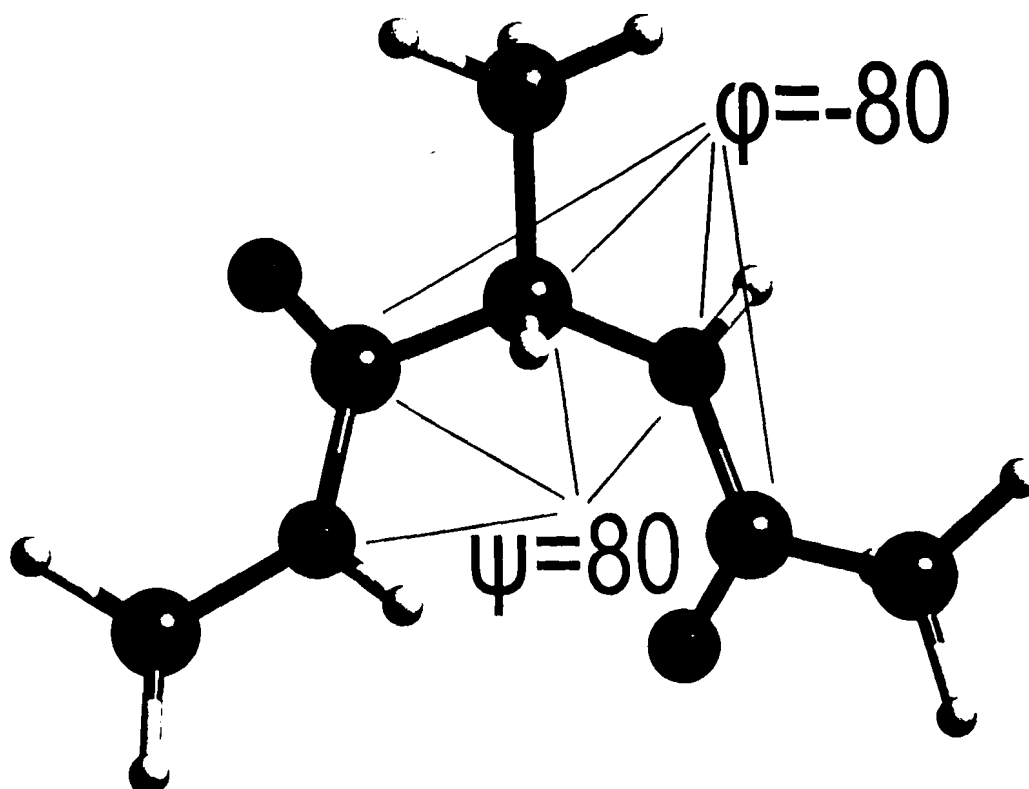
4.6. Εφαρμογές

Αν και η κλασική μοριακή δυναμική έχει συμβάλει αρκετά στην κατανόηση της δομής και λειτουργίας συστημάτων βιολογικής σημασίας^{190,191}, το ενοποιημένο περιβάλλον που αναπτύχθηκε προσφέρει νέες δυνατότητες ώστε να μελετηθούν μοριακά συστήματα, όπου οι δυνάμεις (forces) υπολογίζονται από κβαντικούς υπολογισμούς. Για τον λόγο αυτό εκτελέστηκαν υπολογισμοί σε πεπτιδικά συστήματα.

Η αποδοτική λειτουργικότητα του ενοποιημένου συστήματος ελέγχεται μέσω της αξιολόγησης των αποτελεσμάτων από προσομοιώσεις σε πεπτίδια. Το ενοποιημένο σύστημα *ab-initio* μοριακής δυναμικής που αναπτύχθηκε, χρησιμοποιήθηκε για να προσομοιωθούν τα δύο επόμενα πεπτιδικά μόρια:

1. Το διπεπτίδιο αλανίνης Ace-Ala-Nme.
2. Η κυκλό-εξαγλυκίνη -Gly-Gly-Gly-Gly-Gly-Gly-.

Εικόνα 4.2



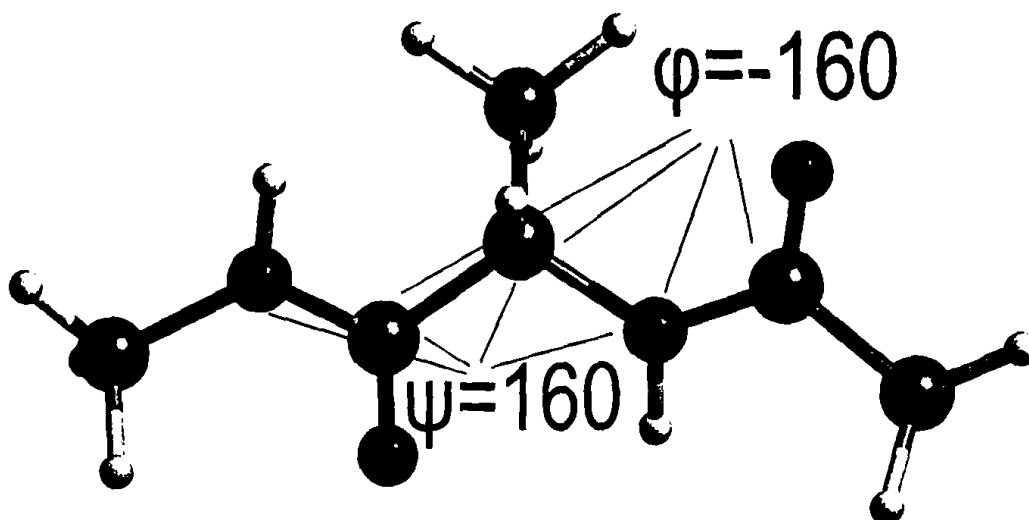
Η C7eq διαμόρφωση του διπεπτιδίου της αλανίνης

4.6.1. Το διπεπτίδιο αλανίνης

Για την προσομοίωση στο διπεπτίδιο της αλανίνης χρησιμοποιήθηκαν δεδομένα από την βιβλιογραφία¹⁹², όπου:

- Έγινε BO ab-initio μοριακή δυναμική με την θεωρία DFT (Density Functional Theory) και το συναρτησιακό (exchange, correlation) των Becke-Perdew¹⁹³.
- Χρησιμοποιήθηκε το σύνολο βάσης DZVP¹⁹⁴.
- Χρησιμοποιήθηκε η μέθοδος ολοκλήρωσης του Verlet.
- Το βήμα ολοκλήρωσης ορίστηκε ως 1.21 fs και η αρχική θερμοκρασία στους 298 K.
- Ο προσομοίωση έγινε σε 70 ps.

Εικόνα 4.3



Η C5 διαμόρφωση του διπεπτιδίου της αλανίνης

Ο διαμορφωσιακός χώρος του διπεπτιδίου περιγράφεται επαρκώς από τις διέδρες γωνίες ϕ και ψ , ενώ οι δύο χαμηλότερες ενεργειακά διαμορφώσεις C7_{eq} (εικόνα 4.2) και C5 (εικόνα 4.3) έχουν τις διέδρες (ϕ, ψ) ίσες με $(-80^\circ, 80^\circ)$ και $(-160^\circ, 160^\circ)$ αντίστοιχα.

4.6.1.1. Υπολογιστικές μέθοδοι προσομοίωσης

Για να εκλεχθεί το ενοποιημένο σύστημα GAMESS-US και TINKER έγινε προσομοίωση στο διπεπτίδιο της αλανίνης με την ημι-εμπειρική μέθοδο AM1 καθώς και με το μοντέλο επιδιалύτωσης SM5.42¹⁹⁵.

Το μοντέλο SM5.42 παρέχεται με το πακέτο GAMESSPLUS¹⁹⁶ το οποίο και διασυνδέθηκε με το πακέτο TINKER.

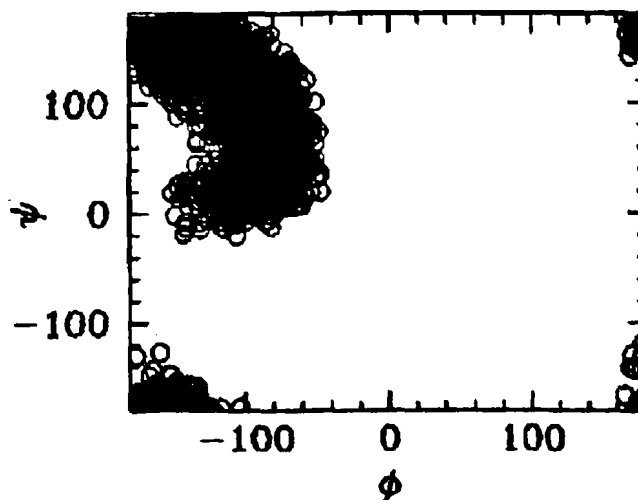
Το GAMESSPLUS είναι μια επέκταση που ενσωματώνεται στο πακέτο GAMESS-US και παρέχει επιπλέον:

- Υπολογισμούς Quantum-Mechanics/Molecular-Mechanics,
- Τα μοντέλα σημειακών φορτίων CM2¹⁹⁷, CM3 και CM4,
- Τα μοντέλα επιδιάλυτωσης SMx (x=5.42, 5.43 και 6).

4.6.1.2. Αποτελέσματα από την προσομοίωση AM1 και SM5.42

Η μοριακή δυναμική στο διπεπίδιο της αλανίνης προσομοιώθηκε σε 70 ps, αντίστοιχα με τον χρόνο της βιβλιογραφικής μελέτης. Μελετήθηκε η μεταβολή των διεδρων γωνιών ϕ και ψ στον χρόνο και τα αποτελέσματα παρουσιάζονται στις εικόνες 4.4 και 4.5.

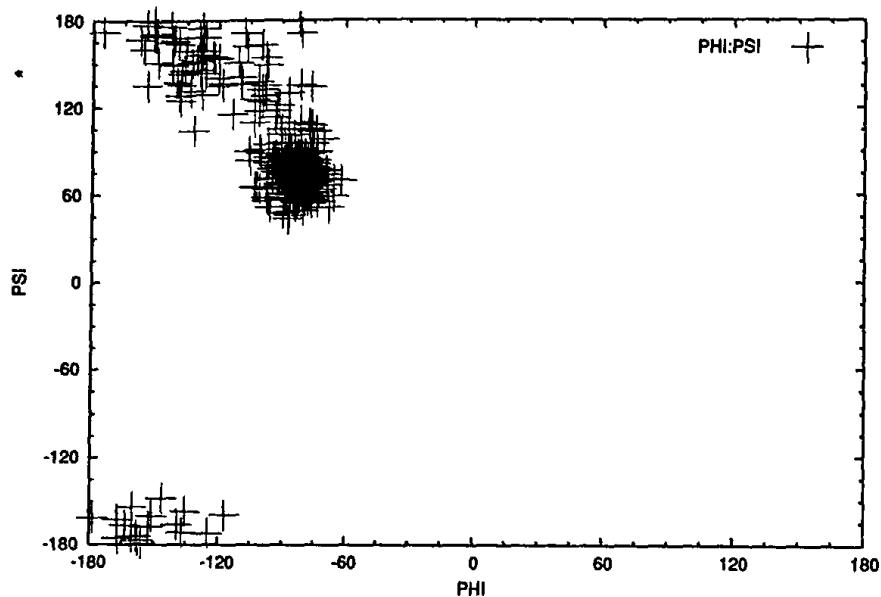
Εικόνα 4.4



Ο διομορφωσιακός χώρος των διεδρων γωνιών από την βιβλιογραφία

Στην συνέχεια η προσομοίωση επαναλήφθηκε προσθέτοντας το μοντέλο επιδιάλυτωσης SM5.42. Οι μεταβολή των διεδρων γωνιών ϕ και ψ με το μοντέλο SM5.42 παρουσιάζεται στην εικόνα 4.5.

Εικόνα 4.5



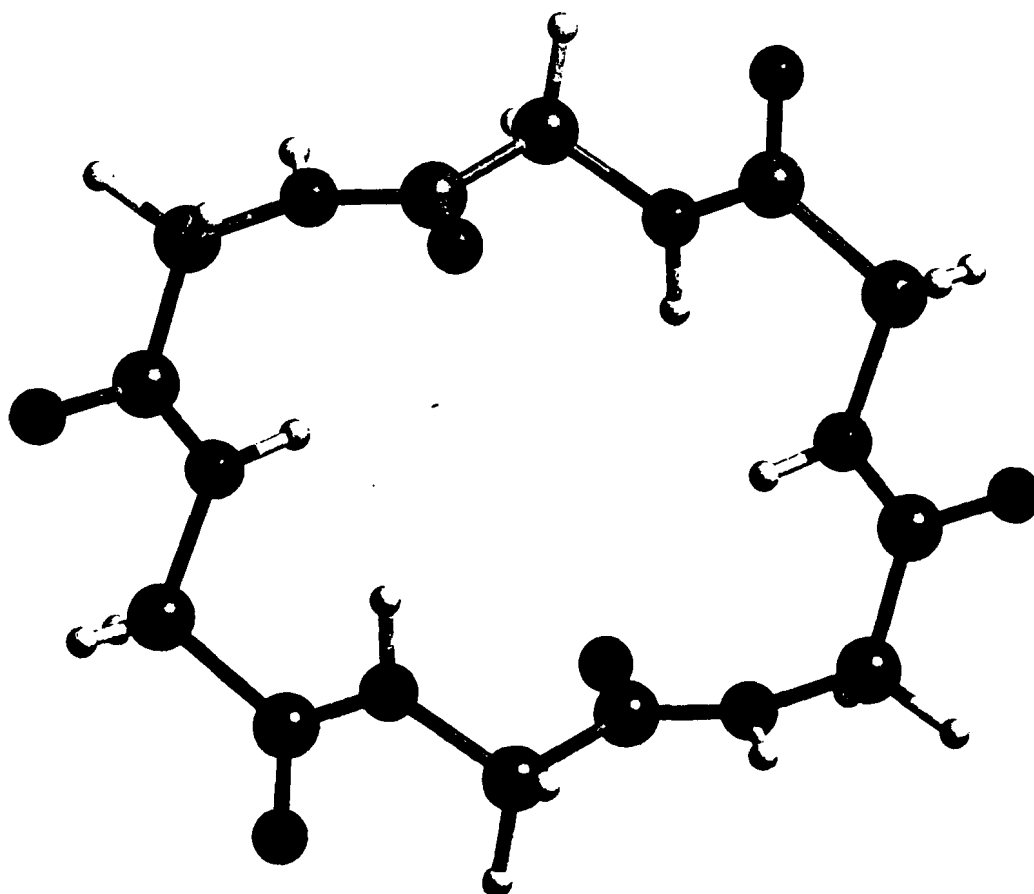
Ο διαμορφωσιακός χώρος (μεταβολή των γωνιών ϕ και ψ) σε χρόνο προσομοίωσης 80 ps με AM1/SM5.42

Στην εικόνα 4.5 φαίνεται ο διαμορφωσιακός χώρος (μεταβολή των διεδρων γωνιών ϕ και ψ) οι οποίες είναι σε πολύ καλή συμφωνία με την βιβλιογραφική υπολογιστική προσομοίωση που αναπτύχθηκε σε επίπεδο θεωρίας DFT.

4.6.2. Η κυκλό-εξαγλυκίνη

Η κυκλό εξαγλυκίνη χρησιμοποιήθηκε ως μοντέλο κυκλικών πεπτιδίων, τα οποία αποτελούν μία σημαντική κατηγορία συστημάτων που χρησιμοποιούνται ως αντιβιοτικά αλλά και γιατί παρέχουν ευκολία στη σύνθεση νέων φαρμάκων¹⁹⁸. Η κυκλό-εξαγλυκίνη και γενικά τα κυκλικά πεπτίδια είναι άκαμπτα¹⁹⁹ μόρια και αυτό που αναμένεται από μια προσομοίωση μοριακής δυναμικής είναι να μην μεταβληθεί το RMSD σημαντικά. Το 1955 η κυκλό-εξαγλυκίνη απομονώθηκε σε κρυσταλλική μορφή²⁰⁰.

Εικόνα 4.6



Η διαμόρφωση της κυκλό-εξαγλυκίνης

Στην προσομοίωση μοριακής δυναμικής που έγινε με το ενοποιημένο σύστημα GAMESS-US και TINKER, επιλέχθηκε η ημι-εμπειρική μέθοδος AM1 με το μοντέλο SM5.42.

Η διάρκεια της προσομοίωσης ορίστηκε στα 150 ps και το βήμα ολοκλήρωσης στα 1 fs. Εξετάστηκε η τροχιά (trajectory) των βαρέων ατόμων (C, O, N) μέσω της μεταβολής του RMSD ως προς τις αρχικές θέσεις των ατόμων του μορίου κοντά στο ελάχιστο (πιο σταθερή διαμόρφωση).

Το RMSD υπολογίζεται ως εξής:

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N |x_i - x_{0i}|^2}$$

όπου N είναι τα άτομα του μορίου,

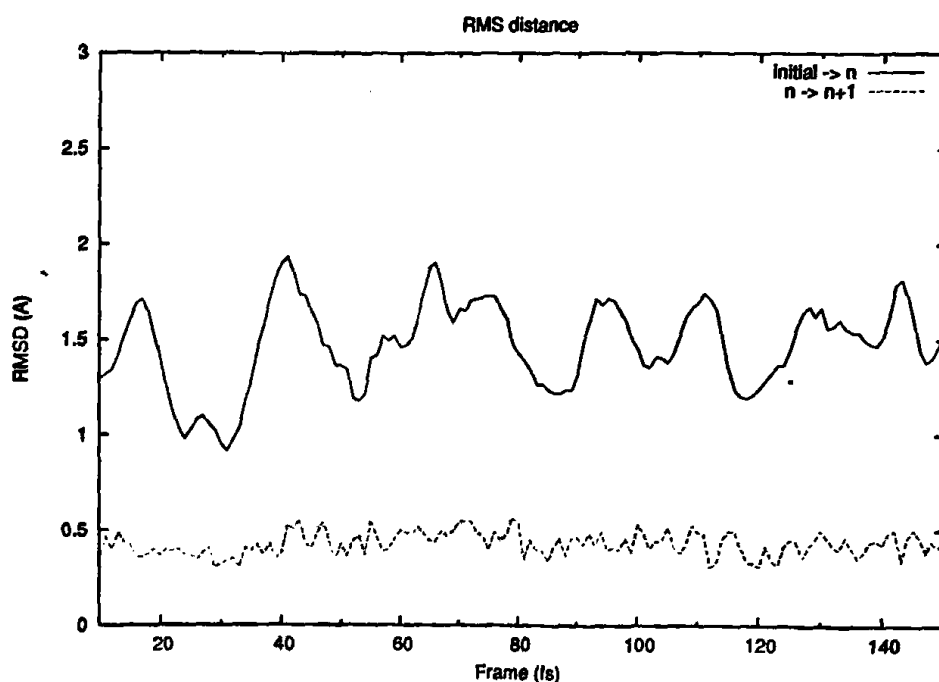
x_i η θέση των ατόμων την τρέχουσα χρονική στιγμή και

x_{0i} η θέση των ατόμων στην αρχή της προσομοίωσης.

Στην εικόνα 4.7 φαίνεται η μεταβολή του RMSD (Å) της κυκλό-εξαγλυκίνης για τα 150 ps της ab initio μοριακής δυναμικής.

Τα αποτελέσματα επιβεβαιώνουν το γεγονός της «ακαμψίας» του κυκλικού εξαπεπτιδίου όπως αναφέρεται και στην βιβλιογραφία μέσω του μικρού γενικά RMSD που παρουσιάζει η προσομοίωση.

Εικόνα 4.7



Η μεταβολή του RMSD σε 150 ps με το μοντέλο AM1-SM5.42



Η πάνω γραμμή (κόκκινη) της εικόνας 4.7 εκφράζει το RMSD της αρχικής διαμόρφωσης ως προς τις επόμενες διαμορφώσεις, ενώ η κάτω γραμμή (πράσινη) εκφράζει το RMSD κάθε διαμόρφωσης με την επόμενη της.

4.6.3. Εκτέλεση προσομοιώσεων με το σύστημα GAMESS-US και TINKER

Για την εκτέλεση μιας προσομοίωσης με τον ενοποιημένο σύστημα απαιτούνται τα επόμενα βήματα:

- Δημιουργία αρχείου εισόδου .inp για το GAMESS-US.
- Δημιουργία αρχείου τύπου .xyz για το TINKER.
- Δημιουργία αρχείου τύπου .key για το TINKER.
- Δημιουργία αρχείου τύπου .run για το TINKER.
- Το script αρχείο «run-md» το οποίο παρέχεται με το περιβάλλον διασύνδεσης.

Στο παράρτημα I δίνονται αυτά τα αρχεία για την εκτέλεση της ab-initio μοριακής δυναμικής του διπεπτιδίου της αλανίνης. Εφόσον δημιουργηθούν τα παραπάνω αρχεία, η προσομοίωση ξεκινά με την εντολή:

```
./run-md acealanme.inp >& acealanmelog.log &
```

Η παράλληλη εκτέλεση του GAMESS-US σε πολλούς επεξεργαστές ή σε cluster υπολογιστών επιταχύνει σημαντικά τον υπολογισμό των δυνάμεων και συνολικά την διαδικασία της ab-initio μοριακής δυναμικής. Η παράλληλη εκτέλεση παρέχει την δυνατότητα ακριβέστερων υπολογιστικών προσομοιώσεων μοριακής δυναμικής, με μεγαλύτερες και καλύτερες κυματοσυναρτήσεις.

4.6.4. Προσομοίωση σε μόρια βιολογικής σημασίας

Το ενοποιημένο σύστημα GAMESS-US και TINKER που αναπτύχθηκε και περιγράφηκε σε αυτό το κεφάλαιο θα εφαρμοστεί για την μελέτη και υπολογιστική προσομοίωση με BO μοριακή δυναμική του πεπτιδικού μοντέλου των μιμητικών σερινοπρωτεάσων που αποτελούν ομάδα πρωτεολυτικών ένζυμων. Η μοντελοποιημένη κατασκευή τους έγινε με το πρόγραμμα Lystas²⁰¹. Το Lystas λειτουργεί ως επέκταση του πακέτου TINKER και έχει ως στόχο την κατασκευή πεπτιδο-μιμητικών μορίων τα οποία πληρούν συγκεκριμένη δομή.



4.7. Συμπεράσματα

Η ανάπτυξη νέων εργαλείων ab-initio μοριακής δυναμικής στοχεύει στην προσομοίωση βιολογικών μορίων, χρησιμοποιώντας όλη την διαθέσιμη υπολογιστική ισχύ με παράλληλους αλγόριθμους καθώς και νέους τρόπους προσέγγισης της μοριακής δυναμικής. Ένα από τα πλεονεκτήματα της προσομοίωσης με μεθόδους ab-initio μοριακής δυναμικής είναι η λεπτομερής περιγραφή των δυναμικών ιδιοτήτων των μοριακών συστημάτων.

Στο κεφάλαιο αυτό περιγράφηκε η δημιουργία ενός νέου ευέλικτου συστήματος ab-initio προσομοιώσεων μοριακής δυναμικής με την ανάπτυξη κώδικα επέκτασης και ενοποίησης των πακέτων TINKER και GAMESS-US. Το σύστημα εκμεταλλεύεται πλήρως την παραλληλοποίηση του πακέτου GAMESS-US δίνοντας την δυνατότητα για την εφαρμογή μεγάλων συνόλων βάσεων και κυματοσυναρτήσεων στην μοριακή δυναμική προσομοίωση.

Η βιβλιοθήκη ανάπτυξης και επέκτασης ρουτινών του πακέτου TINKER χρησιμοποιήθηκε για την ανάπτυξη των νέων ρουτινών για το βήμα της ολοκλήρωσης. Το ενοποιημένο σύστημα δίνει την δυνατότητα για τον επιτυχή υπολογισμό της τροχιάς (trajectory) ενός μοριακού συστήματος.

Η αξιοπιστία του συστήματος, που μπορεί να εκτελεστεί στα λειτουργικά συστήματα UNIX και Linux, μετρήθηκε από την επίτευξη υποδειγματικών προσομοιώσεων του διεπτιδίου της αλανίνης και της κυκλό-εξαγλυκίνης. Τα αποτελέσματα επαληθεύτηκαν επιτυχώς με τη βιβλιογραφία.



Συμπεράσματα της διατριβής

Σε σύγκριση με όλα τα διαθέσιμα συστήματα μοριακής προσομοίωσης και μοντελοποίησης, το ενοποιημένο περιβάλλον MERGAM είναι το μόνο που υποστηρίζει την αυτοματοποιημένη μελέτη μιας περιοχής της επιφάνειας διαμοριακής ενέργειας, όπου ένα μοριακό σύστημα υπόκειται σε κίνηση που καθορίζεται από μαθηματικές εξισώσεις. Επίσης, είναι το μοναδικό σύστημα που παρέχει την δική του γλώσσα προγραμματισμού MCL για την ανάπτυξη ευφύων στρατηγικών διαδικασιών γεωμετρικής αριστοποίησης μοριακών συστημάτων, προσθέτοντας επιπλέον δυνατότητες και ευελιξία στο ευρέως χρησιμοποιούμενο πακέτο GAMESS-US.

Το MERGAM προσφέρει αυτόματη CP-διορθωμένη γεωμετρική βελτιστοποίηση στην επιφάνεια της διαμοριακής ενέργειας αλληλεπίδρασης, διμερών μοριακών συστημάτων, με χρήση πρώτων αναλυτικών παραγώγων. Η δυνατότητα αυτή προγραμματίστηκε και αξιολογήθηκε επιτυχώς στο ενοποιημένο σύστημα.

Για πρώτη φορά έγινε η αυτοματοποιημένη ανακατασκευή και διερεύνηση μια μεγάλης περιοχής της διαμοριακής ενέργειας αλληλεπίδρασης του διμερούς $C_6H_6-H_2O$ με μορφή στρατηγικής διαδικασίας, στην γλώσσα MCL του MERLIN. Η ίδια παραλληλοποιημένη υπολογιστική διαδικασία επαναλήφθηκε με τον πιο αποδοτικό τρόπο σε cluster υπολογιστών, αξιοποιώντας πλήρως τον ενοποιημένο σύστημα.

Το MERGAM είναι ένα εργαλείο που θα εφαρμοστεί για περαιτέρω υπολογιστικές μελέτες που αφορούν πολύπλοκα διαμοριακά προβλήματα τα οποία σχετίζονται με τα γεωμετρικά και ενεργειακά χαρακτηριστικά ασθενών δεσμών υδρογόνου και αλληλεπιδράσεων van der Waals.

Νέες υπολογιστικές διαδικασίες για την μελέτη ασθενών διαμοριακών αλληλεπιδράσεων τύπου OH...π σε ετεροδιμερή μοριακά συστήματα, μέσω ab-initio μεθόδων που στηρίζονται σε πολύ-οριζουσιακές κυματοσυναρτήσεις τύπου CASSCF/NEVPT2, αναμένεται να συμβάλουν θετικά στην ακριβή μοντελοποίηση μεγαλύτερων μοριακών συστημάτων καθώς και στην ανάπτυξη νέων υπολογιστικών τεχνικών και μεθόδων.

Η ιδιότητα της «αυστηρής» size-consistency που παρέχει η μέθοδος NEVPT2 αποτελεί ένα πολύ μεγάλο πλεονέκτημα για την εφαρμογή της υπολογιστικής διαδικασίας που αναπτύχθηκε στα πλαίσια της παρούσας διατριβής για τον υπολογισμό των διαμοριακών ενεργειών αλληλεπίδρασης σε ετεροδιμερή συστήματα. Ο πολύ-οριζουσιακός χαρακτήρας των CP-διορθωμένων OH...π διαμοριακών



αλληλεπιδράσεων, που παίζουν πολύ σημαντικό ρόλο στην διευθέτηση μεγάλων μοριακών συστημάτων βιολογικής σημασίας, αναμένεται να εμπλουτιστεί περαιτέρω με νέες αξιόπιστες μεθόδους που στοχεύουν σε αυτοματοποιημένους υψηλής ακρίβειας υπολογισμούς μοριακής προσομοίωσης. Τα πειραματικά αποτελέσματα, για όλες τις περιπτώσεις που εφαρμόστηκε η θεωρία NEVPT2, σε σύγκριση με τα αποτελέσματα της υπολογιστικής διαδικασίας που αναπτύχθηκε είναι σε πολύ καλή συμφωνία.

Η ανάπτυξη του νέου εργαλείου *ab-initio* μοριακής δυναμικής TINGAM στοχεύει στην προσομοίωση βιολογικών μορίων, χρησιμοποιώντας όλη την διαθέσιμη υπολογιστική ισχύ με παράλληλους αλγόριθμους καθώς και νέους τρόπους προσέγγισης των απ' ευθείας (*on-the-fly*) υπολογισμών της δυναμικής ενέργειας μοριακών συστημάτων. Το σύστημα TINGAM που αναπτύχθηκε με την ενοποίηση των πακέτων TINKER και GAMESS-US εκμεταλλεύεται πλήρως την παραλληλοποίηση του πακέτου GAMESS-US δίνοντας την δυνατότητα για την εφαρμογή μεγάλων συνόλων βάσεων και κυματοσυναρτήσεων στην μοριακή δυναμική προσομοίωση. Η επίτευξη υποδειγματικών προσομοιώσεων του διπεπτιδίου της αλανίνης και της κυκλό-εξαγλυκίνης επαλήθευσε με επιτυχία την αξιοπιστία του συστήματος. Το TINGAM εφαρμόστηκε στην μελέτη των μιμητικών σερινοπρωτεάσων σε συνδυασμό με το εργαλείο LYSTAS που αποτελεί επέκταση του TINKER και έχει ως στόχο την κατασκευή πεπτιδο-μιμητικών μορίων τα οποία πληρούν συγκεκριμένη δομή.

Η τεχνική των συναρτήσεων «wrappers» και «primitives» αποτελεί έναν αξιόπιστο και αποδοτικό τρόπο για την διασύνδεση επιστημονικού λογισμικού προσομοιώσεων, όταν χρειάζεται ενοποίηση και επέκταση υπολογιστικών συστημάτων γραμμένα σε διαφορετικές γλώσσες προγραμματισμού.

Η μεθοδολογία και τεχνολογία που χρησιμοποιήθηκε για την ανάπτυξη των συστημάτων MERGAM και TINGAM είναι δυνατόν να χρησιμοποιηθεί για την διασύνδεση και άλλων πακέτων επιστημονικής προσομοίωσης, παρέχοντας νέες δυνατότητες μοριακής μοντελοποίησης στην ερευνητική και ακαδημαϊκή κοινότητα.

Η αύξηση των υπολογιστικών πόρων που ακολουθεί τον νόμο του Moore (2πλάσια υπολογιστική ισχύ ανά 2 χρόνια) αναμένεται, τα επόμενα χρόνια, να εδραιώσει τις *ab-initio* μοριακές προσομοιώσεις ως κανόνα: α) συνεισφέροντας σε πρωτοποριακά νανοτεχνολογικά επιτεύγματα, και β) κάνοντας πράξη στους καθημερινούς υπολογισμούς προσομοίωσης την τάξη μεγέθους petascale.



ΠΑΡΑΡΤΗΜΑ Α

Η περιγραφή των ρουτίνων C-primitive του συστήματος MERGAM. (Η περιγραφή των συναρτήσεων είναι στην αγγλική γλώσσα για λόγους ορολογίας)

a. `int NewPoint(double *xyz)`

Purpose: Avoids energy or gradient recomputations by checking whether the input geometry has changed since the last call.

Input: `xyz` is a pointer to the atomic coordinates.

Returns: 1 if the energy or gradient is already available by a previous computation at the given coordinates, 0 otherwise.

b. `void SetCharge(int st, int hw, int m)`

Purpose: This function is related to CP-corrected computations. It sets the appropriate nuclear charges to the atoms that constitute a monomer.

Input: `st` is the first atom of a monomer, `hw` is the total number of atoms in the monomer, and `m` is a flag to request monomer computation with dummy (`m=0`) or "ghost" (`m=-1`) atoms.

c. `void GamessRun(double *xyz)`

Purpose: Creates a GAMESS-US input file with the given atomic coordinates and issues a system call to execute the job. The job type is specified according to the global variable `RunType` which assumes the values 0,1, 2 for a job with `RUNTYP=ENERGY`, `GRADIENT` or `HESSIAN` respectively.

Input: `xyz` is a pointer to the atomic coordinates.

d. `void ReadDictionary(void)`

Purpose: Upon completion of a GAMESS-US run, it retrieves the logical record numbers from the dictionary file.

e. `void StoreOrbitals(void)`

Purpose: Stores the molecular orbitals in a disk file in order to be used as an initial prediction in the next GAMESS-US run. This is an implicit way to save computational time and speed up convergence.

f. `void GamessProperty(double *P, enum IrunType RType)`

Purpose: Retrieves the total energy, gradient vector or Hessian matrix from the GAMESS-US dictionary file, according to the value of `RType`.

Input: `P` is a pointer to a memory location where the requested values will be stored. `RType` specifies the requested property as described in the `GamessRun`



function. The value of RType must be less or equal to the value used at the last GamessRun invocation.

g. `void ComputeCPCGrad(double GRD[5][MXATOMS])`

Purpose: This function is related to CP-corrected computations. It provides the CP-corrected gradient vector by evaluating the five terms of equation (4).

Input: GRD is a pointer to a two dimensional array where the computed gradient vectors will be stored.



ΠΑΡΑΡΤΗΜΑ Β

Η περιγραφή των ρουτίνων C-wrapper του συστήματος MERGAM. (Η περιγραφή των συναρτήσεων είναι στην αγγλική γλώσσα για λόγους ορολογίας).

a. `void GMS_INPUT(int *n)`

Purpose: This function is called only once from the main program in order to read the user supplied GAMES-US input file and initialize the MERGAM interface.

Input: `n` is a pointer to the total number of optimization parameters.

b. `void GMS_E(double *xyz, double *en)`

Purpose: Requests a GAMESS-US run to evaluate the total electronic energy. In the case of CP-corrected optimizations it calculates the CP-corrected electronic energy by performing 5 different GAMESS-US runs at the given coordinates, according to equation (2).

Input: `xyz` is a pointer to the atomic coordinates. `en` is a pointer to the total energy.

c. `void GMS_G(double *xyz, double *grd)`

Purpose: Requests a GAMESS-US run to evaluate the gradient vector. In the case of CP-corrected optimizations it calculates the CP-corrected gradient vector by performing 5 different GAMESS-US runs, according to equation (4).

Input: `xyz` is a pointer to the atomic coordinates. `grd` is a pointer to the gradient vector.

d. `void GMS_H(double *xyz, double *hess)`

Purpose: Requests a GAMESS-US run to evaluate the Hessian matrix.

Input: `xyz` is a pointer to the atomic coordinates. `hess` is a pointer to the Hessian matrix.

e. `void GMS_OPTYP(void)`

Purpose: Sets the global flag GRADOPT that controls whether a job with RUNTYP=GRADIENT will be performed, even if only the total electronic energy is required. This will speed up the optimization process when the chosen Merlin optimizer utilizes the gradient vector.

f. `void GMS_CP(double *xyz, int *st, int *hw, double *en, int *typ)`

Purpose: This function is related to CP-corrected computations. It calculates



the total energy of a monomer using dummy or “ghost” atoms.

Input: `xyz` is a pointer to the atomic coordinates. `st` is a pointer to the first atom of the monomer. `hw` is a pointer to the total number of atoms in the monomer. `typ` is a pointer to a flag requesting computation with dummy (`typ=0`) or “ghost” (`typ=-1`) atoms. `en` is a pointer to the computed energy of the monomer.

g. `void GMS_CPCOPT(int *p1, int *p2, int *p3, int *p4)`

Purpose: This function is related to CP-corrected geometry optimizations. It activates the global flag `GCPCOPT` that controls the CP-corrected energy and gradient vector evaluations.

Input: `p1`, `p3` are pointers to the first atom of each monomer, while `p2`, `p4` are pointers to their respective number of atoms.



ΠΑΡΑΡΤΗΜΑ Γ**Εγκατάσταση του περιβάλλοντος MERGAM**

Η επόμενη διαδικασία πρέπει να εφαρμοστεί για την εγκατάσταση του περιβάλλοντος MERGAM. Η διαδικασία προϋποθέτει ότι τα πακέτα GAMESS-US και Merlin-3.1.1 έχουν ήδη εγκατασταθεί.

- a) Το πακέτο MERGAM όπως διανέμεται πρέπει να αποσυμπίεστεί.
- b) Οι παράμετροι εγκατάστασης πρέπει να ρυθμιστούν μέσα στο αρχείο "Makefile.inc".
- c) Ο εκτελέσιμος κώδικας παράγεται με την εντολή "make" στο command-line κέλυφος του λειτουργικού συστήματος ακολουθούμενη από την εντολή "make install".
- d) Το αρχείο εκτέλεσης του GAMESS-US "run-gms" πρέπει να τροποποιηθεί σύμφωνα με τις οδηγίες που παρέχονται στο αρχείο της διανομής "INSTALL.TXT".

Θεωρώντας ότι το σύστημα έχει εγκατασταθεί στον κατάλογο '/usr/local/mergam', το περιβάλλον MERGAM τίθεται σε λειτουργία εκτελώντας την εντολή:

```
/usr/local/mergam/bin/run-mergam filename.inp
```

όπου η παράμετρος είναι ένα αρχείο εισόδου του GAMESS-US και πρέπει πάντα να παρέχεται.

Μετά την εκτέλεση του περιβάλλοντος όλα τα δεδομένα εξόδου από το σύστημα βρίσκονται στο αρχείο "filename.log".

Η μεταγλώττιση κώδικα MCL γίνεται εκτελώντας την εντολή "run-mcl":

```
/usr/local/mergam/bin/run-mcl
```

Για παράδειγμα η πρώτη εφαρμογή του κεφαλαίου 2 εκτελείται ως εξής:

```
run-mergam hfw.inp < autorun >& out.log &
```

και όλα τα απαραίτητα αρχεία βρίσκονται στον κατάλογο '/usr/local/mergam/examples/example1' της διανομής του συστήματος MERGAM.



CPC Panel

Ind Keyword	Description	Value	Allowed values
1) MAF	Monomer A First Atom	1	Any int >= 1
2) MAN	Monomer A Number of Atoms	12	Any int >= 1
3) MBF	Monomer B First Atom	13	Any int >= 1
4) MBN	Monomer B Number of Atoms	3	Any int >= 1
5) CENERGY	non corr/corr/dbsse/cor...	1	{0,1,2,3}
6) CANCEL	Cancel / Proceed	1	{0,1}

Enter changes:

Interaction Energy:	-1.436 Kcal/mol
CP Corrected Int. Energy:	-1.209 Kcal/mol
DELTA BSSE:	-0.227 Kcal/mol
DELTA BSSE(au):	-0.00036251
Total Energy(au):	-306.73761846
Total Corrected Energy(au):	-306.73725596

\\//\\//\\//\\ Merlin is at your command !!!

STOP

----- Merlin run has ended -----



ΠΑΡΑΡΤΗΜΑ Ε

Ο MCL κώδικας που αναπτύχθηκε ως υπόδειγμα στρατηγικής διαδικασίας για την μελέτη περιοχής της υπερεπιφάνειας του διμερούς $C_6H_6-H_2O$.

```

program
var -Oz;H1z;H2z;Ox;Oy;H1x;H1y;H2x;H2y
var k;nx;ny;R0
% request the use of analytic gradients
anal
gmsgrad
% define the intermolecular separation distance at 3.5 Angstrom
R0=3.5
% Loop across the points of the grid plane which is identical with the
% benzene plane
loop nx from -3 to 0.1 by 1
  loop ny from -3 to 0.1 by 1
    % The keyword noeval forces MERLIN to not perform an objective function calculation
    on a
      % new point. This implies that the init command will update only the new point
      variables.
        noeval
        init (npl=1)
        % Store to Oz the oxygen atom z coordinate.
        Oz=x[39]
        % Translate the z coordinates of water's hydrogen atoms to new position
        H1z=x[42]-Oz+R0
        H2z=x[45]-Oz+R0
        % Translate the z coordinate of the oxygen atom to new position.
        % the whole water molecule is now on R0 distance above the benzene plane.
        Oz=R0
        % Store to Ox and Oy the x,y coordinates of oxygen atom.
        Ox=x[37]
        Oy=x[38]
        % Translate the x,y coordinates of water's hydrogen atoms on the nodes of the grid.
        H1x=x[40]-Ox+nx
        H1y=x[41]-Oy+ny
        H2x=x[43]-Ox+nx
        H2y=x[44]-Oy+ny
        % Translate the x,y coordinates of the oxygen atom to new grid position.
        Ox=nx
        Oy=ny
        % As the water molecules lies on the new nx,ny,R0 position we perform a single point
        calculation.
        point
        (x.37=Ox;x.38=Oy;x.39=Oz;x.40=H1x;x.41=H1y;x.42=H1z;x.43=H2x;x.44=H2y;x.45=H2z)
        evaluate
        % Store the geometric and energetic values in file STORE
        memo
        % constrain the benzene molecule. All variables are held fix.
        loop k from 1 to 36 by 1
          fix (x.k)
        -end loop
        % constrain the oxygen atom of water molecule and do not allow any movement.
        fix (x.37;x.38;x.39)
        % call the TOLMIN optimization algorithm with termination criteria the ACC
        % to be less than 0.001
        tolmin (noc=60;acc=1.e-3)
        % Store the new geometry to file STORE
        memo
        % Call the CPC plugin to perform CP-correct energy computations and
        % return the CP-corrected interaction energy as the value of the objective function.
        cpc (maf=1;man=12;mbf=13;mbn=3;cenergy=3)
        % Store the geometric points and energy to file STORE.
        memo
      end loop
    end loop
  end
end

```



ΠΑΡΑΡΤΗΜΑ ΣΤ

```

/*****
*   MERGAM: A programmable optimization environment using the GAMESS-US
*           and MERLIN/MCL packages.
*
*   F.G. Kalatzis, Dept. of Chemistry, Univ. of Ioannina, Greece.
*   Email: tkalatz@cc.uoi.gr
*
*   D.G. Papageorgiou, Dept. of Materials Science and Engineering,
*           Univ. of Ioannina, Greece
*   Email: dpapageo@cc.uoi.gr
*
*   I.N. Demetropoulos, Dept. of Chemistry, Univ. of Ioannina, Greece.
*   Email: idimitr@cc.uoi.gr
*
*****/
*   File: gms_api.c
*   This file contains all MERGAM interface code.
*****/

#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <dirent.h>
#include <math.h>
#include "mer_ext.h"

#if !defined MXATOMS
#   error "MXATOMS was not defined."
#endif

#if !defined RECLEN
#   error "RECLEN was not defined."
#endif

#define DSIZE sizeof(double)
#define LSIZE sizeof(long)

/* Available RUNTYPES */
enum IRunType {ENERGY,GRADIENT,HESSIAN};

/* Function Prototypes */
void ToUpper(char *,char *);
void ReadEnvVariable(char *,char *);
int NewPoint(double *);
void ComputeCPCGrad(double[][MXATOMS]);
void StoreOrbitals(void);
void SetCharge(int,int,int);
void ReadDictionary(void);
void GamessProperty(double *,enum IRunType);
void GamessRun(double *);

/* Global Variables */
long int ioda[400]; /* The GAMESS-US dictionary file entries */
/* Filenames used by MERGAM interface */
char InputFile[256],DictFile[256],PunchFile[256],ExecScript[256];
/* Header Buffer contains all GAMESS-US input sequence until $DATA. */
char Header[200][256];
/* Header Lines */
int HLines;
/* NAT=Number of Atoms, NVAR=Number of variables */
int NAT,NVAR;
enum IRunType RunType;
/* Array storing atom type from GAMESS-US input file */
char atomc[MXATOMS][16];
/* Array that holds atoms charges from GAMESS-US input file */
double chargec[MXATOMS],tcharge[MXATOMS];
int MOW; /* Flag To Append Molecular Orbitals */

```



```

int ST1,HW1,ST2,HW2;
static int GCPCOPT=0; /* Flag to request BSSE-FREE optimization */
static int GRADOPT=0; /* Flag to request ENERGY+GRADIENT evaluations */
/* Atomic coordinates local cache */
double PX[MXATOMS],PY[MXATOMS],PZ[MXATOMS];
/* Final Energy,Gradient and Hessian */
double FE,GRDF[MXATOMS],HESSF[MXATOMS][MXATOMS];
. . .

/* Convert Null-Terminated String to an UpperCase Null-Terminated String */
void ToUpper(char *in,char *out)
{
  int i,len;
  len=strlen(in);
  for (i=0;i<len;i++) {
    out[i]=toupper(in[i]);
  }
  out[i]=0;
}

/* Read Environment Variable var */
void ReadEnvVariable(char *var,char *str)
{
  char *p;
  p=getenv(var);
  if (p==NULL) {
    printf("Invalid Environment Variable %s.\nExiting...\n",var);
    exit(1);
  }
  strcpy(str,p);
}

/* Check if the previous point is identical with current point */
int NewPoint(double *xyz)
{
  int i,flag;
  double X[MXATOMS],Y[MXATOMS],Z[MXATOMS];
  for (i=0;i<NAT;i++) {
    X[i]=*xyz;xyz++;
    Y[i]=*xyz;xyz++;
    Z[i]=*xyz;xyz++;
  }
  for (i=0,flag=0;i<NAT;i++) {
    if (X[i]!=PX[i] || Y[i]!=PY[i] || Z[i]!=PZ[i]) {
      flag=1;
      break;
    }
  }
  if (flag==0) return 0;
  return 1;
}

/* Computes the Analytic counterpoise gradient */
void ComputeCPCGrad(double GRD[][MXATOMS])
{
  int i;
  for (i=0;i<NVAR;i++) {
    GRDF[i]=GRD[0][i]-GRD[1][i]-GRD[2][i]+GRD[3][i]+GRD[4][i];
  }
}

/* Read the orbitals from the Gamess-US PUNCH file and save to file orbitals */
void StoreOrbitals()
{
  FILE *fp,*op;
  int i;
  char line[256],uline[256];

```



```

fp=fopen(PunchFile,"r+t");
if (fp==NULL) {
  printf("Punch File was not Found.Exiting...\n");
  exit(1);
}
op=fopen("orbitals","w+t");
for (i=0;;) {
  fgets(line,256,fp);
  if (feof(fp)) break;
  ToUpper(line,uline);
  if (strstr(uline,"$VEC") i=1;
  if (i==1) fprintf(op,"%s",uline);
  if (strstr(uline,"$END") && i==1) i=0;
}
fprintf(op,"\n");
fclose(op);
fclose(fp);
}

/* Set global charge array for CP computations using ghost or dummy atoms */
void SetCharge(int st,int hw,int m)
{
  int i;
  for (i=st-1;i<(st-1+hw);i++) tcharge[i]=chargec[i]*m;
}

/* This functions read the GAMESS-US dictionary file */
void ReadDictionary()
{
  FILE *fp;
  fp=fopen(DictFile,"r+b");
  if (fp==NULL) {
    printf("GAMESS-US Dictionary File Not Exists.\n");
    exit(1);
  }
  fseek(fp,0L,0);
  fread(&ioda[0],LSIZE*400,1,fp);
  fclose(fp);
}

/* Retrieves the computed property from GAMESS-US Dictionary File */
void GamessProperty(double *P,enum IRunType RType)
{
  FILE *fp;
  fp=fopen(DictFile,"r+b");
  if (fp==NULL) {
    printf("GAMESS-US Dictionary File Not Exists.\n");
    exit(1);
  }
  switch (RType) {
    case ENERGY:  fseek(fp,((ioda[2]-1)*DSIZE*RECLEN)+(2*DSIZE),0);
                   fread(P,DSIZE,1,fp);
                   break;
    case GRADIENT: fseek(fp,(ioda[3]-1)*DSIZE*RECLEN,0);
                   fread(P,NVAR*DSIZE,1,fp);
                   break;
    case HESSIAN:  fseek(fp,(ioda[4]-1)*DSIZE*RECLEN,0);
                   fread(P,NVAR*NVAR*DSIZE,1,fp);
                   break;
    default:      printf("Not a valid property requested\n");
                   exit(1);
                   break;
  }
  fclose(fp);
}

```



```

/* The main GAMESS-US execution function.          */
/* Input parameters:                               */
/*   double *xyz      : the input molecular geometry */
/*   enum RunType RType: the type of computation u   */
void GamessRun(double *xyz)
{
  FILE *fp,*op;
  int, i;
  char temp[256],token[256],*p;
  double x,y,z;
  /* Create the new GAMESS-US input file */
  op=fopen(InputFile,"w+t");
  if_(op==NULL) {
    printf("Error Creating Input File...\nExiting...\n");
    exit(1);
  }
  /* Read the header file and write the new input with the selected RUNTYP */
  for(i=0;i<HLines;i++) {
    strcpy(temp,Header[i]);
    p=strtok(temp," ");
    while(p) {
      ToUpper(p,token);
      if (strstr(token,"RUNTYP=")) {
        if (RunType==ENERGY)   fprintf(op," RUNTYP=ENERGY");
        if (RunType==GRADIENT) fprintf(op," RUNTYP=GRADIENT");
        if (RunType==HESSIAN)  fprintf(op," RUNTYP=HESSIAN");
        p=strtok(NULL," ");
        continue;
      }
      /* Add GUESS=MOREAD to input file so as to use orbitals from previous
run.*/
      if (strstr(token,"GUESS=") && MOW ) {
        fprintf(op," GUESS=MOREAD");
        p=strtok(NULL," ");
        continue;
      }
      fprintf(op," %s",p);
      p=strtok(NULL," ");
    }
  }
  /* end header */
  /* Append the geometry from X,Y,Z arrays to the new input file. */
  for (i=0;i<NAT;i++) {
    x=*xyz;xyz++;
    y=*xyz;xyz++;
    z=*xyz;xyz++;
    fprintf(op,"%s %1f %20.10f %20.10f %20.10f\n",atmc[i],tcharge[i],x,y,z);
  }
  fprintf(op," $END\n");
  /* append the orbitals to the new input file */
  if (MOW) {
    fp=fopen("orbitals","r+t");
    for (;;) {
      fgets(temp,256,fp);
      if (feof(fp)) break;
      fprintf(op,"%s",temp);
    }
    fclose(fp);
  }
  fflush(op);
  fclose(op);
  /* Execute the new GAMESS-US job */
  system(ExecScript);
  /* Read Dictionary File */
  ReadDictionary();
  /* Restore Nuclear Charges */
  for (i=0;i<NAT;i++) tcharge[i]=chargec[i];
  /* Store the orbitals */

```



```

StoreOrbitals();
}

/*****
/*
/* C Wrapper Code - MERLIN F77 Calls */
/*
/*
/*****
/* Activates ENERGY and GRADIENT computation using a single GAMESS-US run */
void GMS_OPTTYP()
{
  GRADOPT=1;
}

/* Prepares the interface for CP-corrected geometry optimization */
/* The code from CPCOPT plugin */
void GMS_CPCOPT(int *p1,int *p2,int *p3,int *p4)
{
  ST1=*p1;
  HW1=*p2;
  ST2=*p3;
  HW2=*p4;
  GCPCOPT=1;
  /* if ST1 equals 0 switch to normal optimizations */
  if (ST1==0) GCPCOPT=0;
  /* change the previous point to force new computation anyway */
  PX[0]=PX[0]+0.000001;
}

/* This function is called initially from main program to perform */
/* interface initialization. */
void GMS_INPUT(int *n)
{
  FILE *fp,*op;
  int i,j,count,args,len=256;
  char line[256],token[256],*p,fname[256];
  char envr[256],envm[256],envs[256],envp[256];
  double xyz;
  /* Read Environment Variables */
  ReadEnvVariable("GMSINP",fname);
  ReadEnvVariable("GMS_R",envr);
  ReadEnvVariable("GMS_M",envm);
  ReadEnvVariable("GMS_SCRATCH",envs);strcat(envs,"/");
  ReadEnvVariable("GMS_PUNCH",envp);strcat(envp,"/");
  /* Molecular orbitals should not be appended. */
  MOW=0;
  /* Initialize job's filenames */
  strcpy(ExecScript,"./");
  strcat(ExecScript,envm);
  strcpy(InputFile,envr);
  strcat(InputFile,".inp");
  strcpy(DictFile,envs);
  strcat(DictFile,envr);
  strcpy(PunchFile,envp);
  strcat(PunchFile,envr);
  strcat(DictFile,".F10");
  strcat(PunchFile,".dat");
  /* Open initial point temp file */
  /* DUMMY is a default MERLIN file that stores the initial point */
  op=fopen("DUMMY","w+t");
  /* Open GAMESS-US input file */
  fp=fopen(fname,"r+t");
  if (fp==NULL) {
    printf("GameSS-US Input File Not Exists.\n");
    exit(1);
  }
}
for (count=0,HLines=0,NAT=0;);) (

```




```

fgets (line, 256, fp);
if (feof(fp)) break;
if (count>0) count++;
ToUpper (line, token);
if (strstr(token, "$DATA")) count++;
if (count>3) {
  if (strstr(token, "$END")) break;
  p=strtok(line, " ");
  strcpy(atomc[NAT], p);
  p=strtok(NULL, " ");
  chargec[NAT]=atof(p);
  tcharge[NAT]=chargec[NAT];
  while (p) {
    p=strtok(NULL, " ");
    if (p==NULL) break;
    if (strstr(p, ".") {
      xyz=atof(p);
      fprintf(op, "%16.12f\n", xyz);
    }
  }
  NAT++;
}
else strcpy(Header[HLines++], line);
}
fclose (fp);
fclose (op);
NVAR=NAT*3;
*n=NVAR;
}

/*The main call from FUNMIN */
void GMS_E(double *xyz, double *en)
{
  int i;
  double DEN1, DEN2, DEN3, DEN4, DBSSE;
  double GRD[5][MXATOMS];
  if (NewPoint(xyz)) {
    if (GRADOPT) { /* Request Energy and Gradient */
      RunType=GRADIENT;
      GamessRun(xyz);
      GamessProperty(&FE, ENERGY);
      GamessProperty(&GRDF[0], GRADIENT);
    } else {
      RunType=ENERGY;
      GamessRun(xyz);
      GamessProperty(&FE, ENERGY);
    }
  }
  MOW=1; /* Next call does not guess MOS */
  if (GCPCOPT==1) { /*CPC Optimization */
    MOW=0; /* CPC calls alwayes guess MOS */
    if (GRADOPT) {
      GamessProperty(&GRD[0][0], GRADIENT);
      SetCharge (ST1, HW1, -1);
      GamessRun (xyz);
      GamessProperty (&DEN1, ENERGY);
      GamessProperty (&GRD[1][0], GRADIENT);
      SetCharge (ST1, HW1, 0);
      GamessRun (xyz);
      GamessProperty (&DEN3, ENERGY);
      GamessProperty (&GRD[3][0], GRADIENT);
      SetCharge (ST2, HW2, -1);
      GamessRun (xyz);
      GamessProperty (&DEN2, ENERGY);
      GamessProperty (&GRD[2][0], GRADIENT);
      SetCharge (ST2, HW2, 0);
      GamessRun (xyz);
      GamessProperty (&DEN4, ENERGY);
      GamessProperty (&GRD[4][0], GRADIENT);
    }
  }
}

```



```

    ComputeCPCGrad (GRD);
  } else {
    SetCharge (ST1, HW1, -1);
    GamessRun (xyz);
    GamessProperty (&DEN1, ENERGY);
    SetCharge (ST1, HW1, 0);
    GamessRun (xyz);
    GamessProperty (&DEN3, ENERGY);
    SetCharge (ST2, HW2, -1);
    GamessRun (xyz);
    GamessProperty (&DEN2, ENERGY);
    SetCharge (ST2, HW2, 0);
    GamessRun (xyz);
    GamessProperty (&DEN4, ENERGY);
  }
  DBSSE=DEN1-DEN3+DEN2-DEN4;
  FE--DBSSE;
}
}
*en=FE;
for (i=0;i<NVAR;i++) {
  PX[i]=*xyz;xyz++;
  PY[i]=*xyz;xyz++;
  PZ[i]=*xyz;xyz++;
}
}

/* The main call from GRANAL */
void GMS_G(double *xyz,double *grd)
{
  int i;
  double GRD[5][MXATOMS];
  /* If new point or previous run was energy */
  if (NewPoint(xyz) || RunType==ENERGY) {
    RunType=GRADIENT;
    GamessRun (xyz);
    GamessProperty (&FE, ENERGY);
    GamessProperty (&GRDF[0], GRADIENT);
    MOW=1;
    if (GCPCOPT==1) {
      MOW=0;
      GamessProperty (&GRD[0][0], GRADIENT);
      SetCharge (ST1, HW1, -1);
      GamessRun (xyz);
      GamessProperty (&GRD[1][0], GRADIENT);
      SetCharge (ST1, HW1, 0);
      GamessRun (xyz);
      GamessProperty (&GRD[3][0], GRADIENT);
      SetCharge (ST2, HW2, -1);
      GamessRun (xyz);
      GamessProperty (&GRD[2][0], GRADIENT);
      SetCharge (ST2, HW2, 0);
      GamessRun (xyz);
      GamessProperty (&GRD[4][0], GRADIENT);
      ComputeCPCGrad (GRD);
    }
  }
  for (i=0;i<NVAR;i++) {
    *grd=GRDF[i];
    grd++;
  }
}

/* The main call from HANAL */
void GMS_H(double *xyz,double *hess)
{

```



```
int i,j;
RunType=HESSIAN;
GamessRun(xyz);
GamessProperty(&FE,ENERGY);
GamessProperty(&GRDF[0],GRADIENT);
GamessProperty(&HESSF[0][0],HESSIAN);
for (i=0;i<NVAR;i++)
  for (j=0;j<NVAR;j++) {
    *hess=HESSF[i][j];
    *hess++;
  }
}

/* C Wrapper Code that calls the GamessRun function with typ=0 */
/* or 1 to perform single point counterpoise energy computation.*/
/* This function is called from the CPC plugin. */
void GMS_CP(double *xyz,int *st,int *hw,double *en,int *typ)
{
  RunType=ENERGY;
  MOW=0;
  switch (*typ) {
    case 0: SetCharge(*st,*hw,-1);
            GamessRun(xyz);
            break;
    case 1: SetCharge(*st,*hw,0);
            GamessRun(xyz);
            break;
  };
  GamessProperty(en,ENERGY);
}
```



ΠΑΡΑΡΤΗΜΑ Ζ

Ο επόμενος κώδικας παράγει την λίστα με τις διπλές διεγέρσεις που χρησιμοποιούνται για την παραγωγή των οριζουσών σε ένα υπολογισμό με την θεωρία CI.

```
#include <stdio.h>
#include <math.h>
#include <string.h>

int main(int argc, char *argv[])
{
    int i, j, k, l;
    int N, M;
    int oc[1000], uc[1000];
    N=atoi(argv[1]);
    M=atoi(argv[2]);
    printf(" $GCILST\n");
    printf("%d\n", (N*M)+1);
    for (i=0; i<N; i++) {oc[i]=2; printf("%-2d", oc[i]);}
    for (j=0; j<M; j++) {uc[j]=0; printf("%-2d", uc[j]);}
    printf("\n");
    for (i=0; i<N; i++) {
        for (j=0; j<M; j++) {
            for (k=N-1; k>-1; k--) {
                if (k==i) printf("%-2d", 0);
                else printf("%-2d", oc[k]);
            }
            for (k=0; k<M; k++) {
                if (k==j) printf("%-2d", 2);
                else printf("%-2d", uc[k]);
            }
            printf("\n");
        }
    }
    printf(" $END\n");
    return 0;
}
```



ΠΑΡΑΡΤΗΜΑ Η

Ο κώδικας της ρουτίνας «beeman2.f» για την μέθοδο ολοκλήρωσης του Beeman.

```

c
c   "beeman" performs a single molecular dynamics time step
c   by means of a Beeman multistep recursion formula; the
c   actual coefficients are Brooks' "Better Beeman" values
c
c   literature reference:
c
c   B. R. Brooks, "Algorithms for Molecular Dynamics at Constant
c   Temperature and Pressure", DCRT Report, NIH, April 1988
c
c
c   subroutine beeman2(istep,dt,NNN)
c   implicit none
c   include 'sizes.i'
c   include 'atmtyp.i'
c   include 'atoms.i'
c   include 'moldyn.i'
c   include 'shake.i'
c   include 'units.i'
c   include 'usage.i'
c   integer i,j,istep,NNN,ii
c   real*8 XX(NNN),GRAD(NNN)
c   real*8 dt,etot,En
c   real*8 dt_8,dt2_8
c   real*8 eksum,epot
c   real*8 temp,pres
c   real*8 xterm,yterm,zterm
c   real*8 ekin(3,3)
c   real*8 stress(3,3)
c   real*8 xold(maxatm)
c   real*8 yold(maxatm)
c   real*8 zold(maxatm)
c   real*8 derivs(3,maxatm)
c
c
c   set some time values for the dynamics integration
c
c   dt_8 = dt / 8.0d0
c   dt2_8 = dt * dt_8
c
c   make half-step temperature and pressure corrections
c
c   call temper (dt)
c
c   store the current atom positions, then find new atom
c   positions and half-step velocities via Beeman recursion
c
c   do i = 1, n
c     if (use(i)) then
c       xold(i) = x(i)
c       yold(i) = y(i)
c       zold(i) = z(i)
c       xterm = 5.0d0*a(1,i) - aold(1,i)
c       yterm = 5.0d0*a(2,i) - aold(2,i)
c       zterm = 5.0d0*a(3,i) - aold(3,i)
c       x(i) = x(i) + v(1,i)*dt + xterm*dt2_8
c       y(i) = y(i) + v(2,i)*dt + yterm*dt2_8
c       z(i) = z(i) + v(3,i)*dt + zterm*dt2_8
c       v(1,i) = v(1,i) + xterm*dt_8
c       v(2,i) = v(2,i) + yterm*dt_8
c       v(3,i) = v(3,i) + zterm*dt_8
c     end if
c   end do
c
c   get constraint-corrected positions and half-step velocities

```



```

c
  if (use_rattle) call rattle (dt,xold,yold,zold)
c
  get the potential energy and atomic forces
c
  call gradient (epot,derivs)
  ii=1
  do i=1,n
    XX(ii)=x(i)
    ii=ii+1
    XX(ii)=y(i)
    ii=ii+1
    XX(ii)=z(i)
    ii=ii+1
  end do
  call GMS_E(XX,En)
  call GMS_G(XX,GRAD)
  epot=En*627.5095
  ii=1
  do i=1,n
    derivs(1,i)=GRAD(ii)*627.5095
    ii=ii+1
    derivs(2,i)=GRAD(ii)*627.5095
    ii=ii+1
    derivs(3,i)=GRAD(ii)*627.5095
    ii=ii+1
  c
    write(*,*)derivs(1,i),derivs(2,i),derivs(3,i)
  end do
c
  use Newton's second law to get the next accelerations;
  find the full-step velocities using the Beeman recursion
c
  do i = 1, n
    if (use(i)) then
      do j = 1, 3
        aold(j,i) = a(j,i)
        a(j,i) = -convert * derivs(j,i) / mass(i)
        v(j,i) = v(j,i) + (3.0d0*a(j,i)+aold(j,i))*dt_8
      end do
    end if
  end do
c
  find the constraint-corrected full-step velocities
c
  if (use_rattle) call rattle2 (dt)
c
  accumulate the kinetic energy and its outer product
c
  call kinetic (eksum,ekin)
c
  make full-step temperature and pressure corrections
c
  call temper2 (dt,eksum,temp)
  call pressure (dt,ekin,pres,stress)
c
  system energy is sum of kinetic and potential energies
c
  etot = eksum + epot
c
  compute statistics and save trajectory for this step
c
  call mdstat (istep,dt,etot,epot,eksum,temp,pres)
  call mdsave (istep,dt,epot)
  return
  end

```



ΠΑΡΑΡΤΗΜΑ Θ

Ο κώδικας της ρουτίνας «verlet2.f» για την μέθοδο ολοκλήρωσης του Verlet.

```

c
c      "verlet" performs a single molecular dynamics time step
c      by means of the velocity Verlet multistep recursion formula
c
c
c      * *
c      subroutine verlet2(istep,dt,NNN)
c      implicit none
c      include 'sizes.i'
c      include 'atmtyp.i'
c      include 'atoms.i'
c      include 'moldyn.i'
c      include 'shake.i'
c      include 'units.i'
c      include 'usage.i'
c      integer i,j,istep,NNN,ii
c      real*8 XX(NNN),GRAD(NNN)
c      real*8 dt,etot,En
c      real*8 dt_2,dt2_2
c      real*8 eksum,epot
c      real*8 temp,pres
c      real*8 ekin(3,3)
c      real*8 stress(3,3)
c      real*8 xold(maxatm)
c      real*8 yold(maxatm)
c      real*8 zold(maxatm)
c      real*8 derivs(3,maxatm)
c
c
c      set some time values for the dynamics integration
c
c      dt_2 = dt / 2.0d0
c      dt2_2 = dt * dt_2
c
c      make half-step temperature and pressure corrections
c
c      call temper (dt)
c
c      store the current atom positions, then find new atom
c      positions and half-step velocities via Verlet recursion
c
c      do i = 1, n
c         if (use(i)) then
c            xold(i) = x(i)
c            yold(i) = y(i)
c            zold(i) = z(i)
c            x(i) = x(i) + v(1,i)*dt + a(1,i)*dt2_2
c            y(i) = y(i) + v(2,i)*dt + a(2,i)*dt2_2
c            z(i) = z(i) + v(3,i)*dt + a(3,i)*dt2_2
c            do j = 1, 3
c               v(j,i) = v(j,i) + a(j,i)*dt_2
c            end do
c         end if
c      end do
c
c      get constraint-corrected positions and half-step velocities
c
c      if (use_rattle) call rattle (dt,xold,yold,zold)
c
c      get the potential energy and atomic forces
c
c      call gradient (epot,derivs)
c      ii=1
c      do i=1,n
c         XX(ii)=x(i)
c         ii=ii+1

```



```

XX(ii)=y(i)
ii=ii+1
XX(ii)=z(i)
ii=ii+1
end do
call GMS_E(XX,En)
call GMS_G(XX,GRAD)
epot=En*627.5095
ii=1
do i=1,n
  derivs(1,i)=GRAD(ii)*627.5095
  ii=ii+1
  derivs(2,i)=GRAD(ii)*627.5095
  ii=ii+1
  derivs(3,i)=GRAD(ii)*627.5095
  ii=ii+1
c   write(*,*)derivs(1,i),derivs(2,i),derivs(3,i)
end do

c
c   use Newton's second law to get the next accelerations;
c   find the full-step velocities using the Verlet recursion
c
do i = 1, n
  if (use(i)) then
    do j = 1, 3
      a(j,i) = -convert * derivs(j,i) / mass(i)
      v(j,i) = v(j,i) + a(j,i)*dt_2
    end do
  end if
end do

c
c   find the constraint-corrected full-step velocities
c
if (use_rattle) call rattle2 (dt)

c
c   accumulate the kinetic energy and its outer product
c
call kinetic (eksum,ekin)

c
c   make full-step temperature and pressure corrections
c
call temper2 (dt,eksum,temp)
call pressure (dt,ekin,pres,stress)

c
c   system energy is sum of kinetic and potential energies
c
etot = eksum + epot

c
c   compute statistics and save trajectory for this step
c
call mdstat (istep,dt,etot,epot,eksum,temp,pres)
call mdsave (istep,dt,epot)
return
end

```



ΠΑΡΑΡΤΗΜΑ Ι

1. Το αρχείο εισόδου του GAMESS-US «acealanme.inp»:

```

$contrl scftyp=RHF RUNTYP=GRADIENT nzvar=0
        exetyp=run maxit=200 MULT=1 ICHARG=0 units=angs
        inttyp=pople coord=unique mplevl=0 ICUT=9 ITOL=12 $end
$system mwords=46 PARALL=.FALSE. $end
$guess  GUESS=huckel $end
$basis  gbasis=AM1 $end
$mopac  PEPTID=.TRUE. $end
$scf    conv=8 $end
$mp2    nacore=0 cutoff=1.d-9 $end
$data
Comment
C1
C 6.0 0.000000 0.000000 0.000000
C 6.0 0.000000 0.000000 1.510000
O 8.0 1.028938 0.000000 2.165506
H 1.0 1.056270 0.000000 -0.341166
H 1.0 -0.497647 0.931694 -0.341166
H 1.0 -0.497647 -0.931694 -0.341166
N 7.0 -1.236201 0.000000 2.027114
C 6.0 -1.446961 0.000000 3.471822
C 6.0 -2.492263 -1.009926 3.881081
O 8.0 -3.493695 -0.700499 4.505411
H 1.0 -2.025518 0.000000 1.381066
H 1.0 -0.498513 -0.263763 3.984633
C 6.0 -1.934563 1.390255 3.920192
H 1.0 -2.094246 1.388399 5.018645
H 1.0 -2.890701 1.628178 3.409024
H 1.0 -1.169185 2.150967 3.660166
N 7.0 -2.192264 -2.254273 3.484594
C 6.0 -3.080189 -3.373713 3.784664
H 1.0 -1.324917 -2.395866 2.966847
H 1.0 -2.649888 -4.309310 3.370419
H 1.0 -4.075331 -3.189877 3.328599
H 1.0 -3.192771 -3.472852 4.884481
$END

```

2. Το αρχείο εισόδου του TINKER «acealanme.xyz»:

```

22 acealanme
 1 CT 0.000000 0.000000 0.000000 340 2 4 5 6
 2 C 0.000000 0.000000 1.510000 342 1 3 7
 3 O 1.028938 0.000000 2.165506 343 2
 4 HC 1.056270 0.000000 -0.341166 341 1
 5 HC -0.497647 0.931694 -0.341166 341 1
 6 HC -0.497647 -0.931694 -0.341166 341 1
 7 N -1.236201 0.000000 2.027114 7 2 8 11
 8 CT -1.446961 0.000000 3.471822 8 7 9 12 13
 9 C -2.492263 -1.009926 3.881081 9 8 10 17
 0 O -3.493695 -0.700499 4.505411 11 9
11 H -2.025518 0.000000 1.381066 10 7
12 H1 -0.498513 -0.263763 3.984633 12 8
13 CT -1.934563 1.390255 3.920192 13 8 14 15 16
14 HC -2.094246 1.388399 5.018645 14 13
15 HC -2.890701 1.628178 3.409024 14 13
16 HC -1.169185 2.150967 3.660166 14 13
17 N -2.192264 -2.254273 3.484594 346 9 18 19
18 CT -3.080189 -3.373713 3.784664 348 17 20 21 22
19 H -1.324917 -2.395866 2.966847 347 17
20 H1 -2.649888 -4.309310 3.370419 349 18
21 H1 -4.075331 -3.189877 3.328599 349 18
22 H1 -3.192771 -3.472852 4.884481 349 18

```



3. Το αρχείο εισόδου του TINKER «acealanme.key»:

```
PARAMETERS /opt/tinker/params/amber99.prm
INTEGRATE VERLET
VERBOSE
SAVE-VELOCITY
SAVE-FORCE
```

4. Το αρχείο εισόδου του TINKER «acealanme.run»:

```
./dynamicab.x acealanme.xyz 500000 1 1 300 -k acealanme.key
```

5. Το αρχείο εκτέλεσης της προσομοίωσης «run-md»:

```
#!/bin/sh
GMS_SCRATCH=/opt/gamess4/output
GMS_PUNCH=/opt/gamess4/output
MDEXE=./acealanme.run
GMS_EXEC=/opt/gamess4/rgms
GMS_M=mdrun
GMS_R=mer$$
EXECCM=$MDEXE
CP=cp
RM="rm -f"
MV=mv
BASENAME=basename
CAT=cat
if [ $# -ne 1 ]
then
    echo
    echo Usage: `basename $0` gamess_input_file
    exit 1
fi
echo -e "#!/bin/sh" > ./GMS_M
echo -e "$RM $GMS_SCRATCH/$GMS_R.* > /dev/null 2>&1" >> ./GMS_M
echo -e "$MV $GMS_R.inp $GMS_R.inp > /dev/null 2>&1" >> ./GMS_M
echo -e "$RM $GMS_PUNCH/$GMS_R.* > /dev/null 2>&1" >> ./GMS_M
echo -e "$MV $GMS_R.inp $GMS_R.inp > /dev/null 2>&1" >> ./GMS_M
echo -e "$GMS_EXEC $GMS_R > $GMS_R.log 2>&1" >> ./GMS_M
chmod +x ./GMS_M
GMSINP=$1
if [ ! -f $GMSINP ]
then
    echo The GAMESS input file $GMSINP does not exist.
    exit 2
fi
$RM $GMS_R.log > /dev/null 2>&1
export GMSINP
export GMS_R
export GMS_M
export GMS_PUNCH
export GMS_SCRATCH
$EXECCM
$RM $GMS_SCRATCH/$GMS_R.F* > /dev/null 2>&1
$RM orbitals $GMS_M $GMS_R.inp t > /dev/null 2>&1
BASE="$BASENAME $GMSINP.inp"
if [ -f $GMS_R.log ]
then
    $MV $GMS_R.log $BASE.log
fi
```



ΠΑΡΑΡΤΗΜΑ Κ

Μέθοδοι Κβαντικών Υπολογισμών

Μέθοδος	Περιγραφή
Conventional RHF	Restricted Hartree-Fock with the integrals stored on disk
Direct RHF	Restricted Hartree-Fock with the integrals computed as needed
In-core RHF	Restricted Hartree-Fock with the integrals held in memory
Analytical RHF Gradient	Restricted Hartree-Fock first derivatives of the energy Gradient
Analytical RHF Hessian	Restricted Hartree-Fock second derivatives of the energy
Conventional UHF	Unrestricted Hartree-Fock with the integrals stored on disk
Conventional MP2	Second order Møller-Plesset Perturbation theory (disk based)
Direct MP2	Second order Møller-Plesset Perturbation theory (direct)
Analytical MP2 Gradient	Second order Møller-Plesset Perturbation theory first derivatives
Analytical MP2 Hessian	Second order Møller-Plesset Perturbation theory second derivatives
MP4(SDTQ)	Fourth order Møller-Plesset Perturbation theory (disk-based)
SDCI	Singles and doubles configuration interaction (1 ref. configuration)
CCSD	Coupled clusters with singles and doubles
CCSD(T)	Coupled clusters with singles and doubles and perturbative triples
QCISD	Quadratic configuration interaction with singles and doubles
QCISD(T)	Quadratic CI with singles and doubles and perturbative triples
CASSCF	Complete Active Space Self Consistent Field energy evaluation
CAS-CI	CAS-based multireference configuration interaction



ΠΑΡΑΡΤΗΜΑ Μ

Πακέτα υπολογιστικής προσομοίωσης στην χημεία.

Package	Molecular Mechanics	Semi- Empirical	Hartree- Fock	Post- Hartree- Fock methods	Density Functional Theory
COLUMBUS	N	N	Y	Y	N
DALTON	N	N	Y	Y	Y
GAMESS (UK)	N	Y	Y	Y	Y
GAMESS (US)	N	Y	Y	Y	Y
GAUSSIAN	Y	Y	Y	Y	Y
MOLCAS	Y	Y	Y	Y	Y
MOLPRO	N	N	Y	Y	Y
MPQC	N	N	Y	Y	Y
NWChem	Y	N	Y	Y	Y
PSI	N	N	Y	Y	N
Q-Chem	Y	N	Y	Y	Y
TURBOMOLE	Y	N	Y	Y	Y



ΠΑΡΑΡΤΗΜΑ Ν

Α. Σταθερές που χρησιμοποιούνται στην μοριακή κβαντική προσομοίωση.

$$1 \text{ Bohr } (a_0) = 0.5291772083 \text{ \AA}$$

$$1 \text{ Atomic mass unit (amu)} = 1.66053873 \times 10^{-27} \text{ kilograms}$$

$$1 \text{ Electron charge (e)} = 4.803242 \times 10^{10} \text{ ESU} = 1.602176462 \times 10^{-19}$$

Coulombs

$$\text{Planck's constant (} h \text{)} = 6.62606876 \times 10^{-34} \text{ Joule-secs}$$

$$\text{Avogadro's number (} N_A \text{)} = 6.02214199 \times 10^{23}$$

$$1 \text{ Calorie} = 4.184 \text{ Joules}$$

$$1 \text{ Hartree} = 4.3597482 \times 10^{-18} \text{ Joules}$$

$$\text{Speed of light (} c \text{)} = 2.99792458 \times 10^{10} \text{ cm-sec}^{-1}$$

$$\text{Boltzman constant (} k \text{)} = 1.380603 \times 10^{-23} \text{ Joules-degree}^{-1}$$

$$\text{Inverse fine structure constant (} \alpha^{-1} \text{)} = 137.03599976$$

$$\text{Molar volume of ideal gas at 273.15 K} = 0.022413996 \text{ m}^3$$

$$\text{Proton rest mass} = 1.67262158 \times 10^{-27} \text{ kg}$$

$$\text{Electron magnetic moment} = 9.28476362 \times 10^{-24} \text{ J-T}^{-1}$$

$$\text{Free electron g-factor} = 2.002319304386 \text{ (dimensionless)}$$

Β. Συντελεστές μετασχηματισμού μονάδων.

$$\text{Electron mass} = 0.910938 \times 10^{-30} \text{ kg}$$

$$\text{Proton mass} = 1836.1527 \text{ electron mass}$$

$$1 \text{ Atomic mass unit (amu)} = 1822.8880 \text{ electron mass}$$

$$1 \text{ Electron volt (eV)} = 23.06055 \text{ kcal-mol}^{-1}$$

$$1 \text{ Hartree} = 627.5095 \text{ kcal-mol}^{-1} = 27.2114 \text{ eV}$$

$$1 \text{ Bohr-electron} = 2.541746 \text{ Debye} = 2.541746 \times 10^{-18} \text{ esu-cm}$$

$$1 \text{ Debye}^2\text{-angstrom}^2\text{-amu}^{-1} = 42.2561 \text{ km-mol}^{-1} = 5.82573 \times 10^{-3} \text{ cm}^2\text{-atm}^{-1} \text{ at STP}$$

$$1 \text{ Hartree}^{1/2}\text{-Bohr-1-amu}^{-1/2} = 219474.6 \text{ cm}^{-1}$$

$$\text{Electric field: } 1 \text{ au} = 5.142206 \times 10^{11} \text{ V-m}^{-1}$$

$$\text{Electric polarizability: } 1 \text{ au} = 1.648777 \times 10^{-41} \text{ C}^2\text{-m}^2\text{-J}^{-1}$$

$$\text{Dipole moment} = 1 \text{ Bohr-electron} = 8.478352 \times 10^{-30} \text{ C-m}$$



Βιβλιογραφία

1. Intl. Trans. in Op. Res. 13, 85 (2006).
2. J. Reitman, "A Concise History of the Ups and Downs of Simulation".
3. N. Metropolis et al., J. Chem. Phys. 21, 1087 (1953).
4. M. N. Rosenbluth, A. W. Rosenbluth, J. Chem. Phys. 22, 881 (1954).
5. W. W. Wood, F. R. Parker, J. Chem. Phys. 27, 720 (1957).
6. H. Popkie, H. Kistenmacher, and E. Clementi, J. Chem. Phys. 59, 1325 (1973).
7. M. Hloucha, A. K. Sum, S. I. Sandler, J. Chem. Phys., 113, 13, (2000).
8. R. Bukowski, K. Szalewicz, and C. F. Chabalowski, J. Phys. Chem. A 103, 7322 (1999).
9. http://nobelprize.org/nobel_prizes/chemistry/laureates/1998/
10. F. Gygi, Journal of Physics, Conference Series 46, 268 (2006).
11. Thunder is a 23 TFlop platform installed at LLNL in 2004. It comprises 4096 Itanium2 processors in 1024 nodes.
12. A. Gara et al., IBM J. Res. Dev., 49, 195 (2005).
13. <http://cpc.cs.qub.ac.uk/>
14. X. Jiao, G. Zheng, P. A. Alexander, M. T. Campbell, O. S. Lawlor, J. Norris, A. Haselbacher, M. T. Heath, A system integration framework for coupled multiphysics simulations, Engineering with Computers, 22, 293 (2006).
15. E. Sturler, J. Hoeflinger, L. Kale, M. Bhandarkar, Frameworks for Multi-physics Simulation Codes, IFIP Conference Proceedings, Vol. 188, 87, (2000).
16. http://www.csar.uiuc.edu/F_info/AboutCSAR.htm
17. D. Rantzaou, U. Lang, A Scalable Virtual Environment for Large Scale Scientific Data Analysis, Future Generation Computer Systems 14, 215 (1998).
18. T. Breitfeld., S. Kolibal., A. Schreiber, W. Wagner, "Java for controlling and configuring a distributed Turbine Simulation System", Workshop Java for High Performance Network Computing, Southampton (1998).
19. P.T. Homer, R.D. Schlichting, Concurrency Practice and Experience, 6, 27 1 (1994).
20. P.F. Dubois, L. Busby, "Portable, Powerful Fortran Programs", Computers in Physics, 7, No. 1, (1993).
21. <http://paw.web.cern.ch/paw/>
22. T. Steinke, Modern Methods and Algorithms of Quantum Chemistry, J. Grotendorst (Ed.), John von Neumann Institute for Computing, 1, ISBN 3-00-005618-1, 49 (2000).
23. M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, J. Comput. Chem. 14, 1347 (1993).
24. G.D. Fletcher, M.W. Schmidt, Distributed Data Interface – "SPMD" Data Server Model, (1999).
25. Gaussian 98, Revision A.7, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Cli_ ord, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T.



- Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, Gaussian, Inc., Pittsburgh PA, 1998. <http://www.gaussian.com>
26. MPQC (Massively Parallel Quantum Chemistry) Limit Point Systems, Inc., Fremont, CA, 1999. <http://aros.ca.sandia.gov/cljanss/mpqc/>
 27. C.L. Janssen, E.T. Seidl, and M.E. Colvin, Object Oriented Implementation of Parallel Ab Initio Programs, ACS Symposium Series 592, Parallel Computers in Computational Chemistry, T. Mattson, Ed., 1995. <http://aros.ca.sandia.gov/cljanss/sc/>
 28. R.J. Harrison, The TCGMSG Message-Passing Toolkit, Pacific Northwest National Laboratory, version 4.04, (1994).
<http://www.emsl.pnl.gov:2080/docs/nwchem/tcgmsg.html>
 29. <http://dasher.wustl.edu/tinker/summary.html>
 30. S.P. Webb, T. Iordanov, and S. Hammes-Schiffer, *J. Chem. Phys.* 117, 4106 (2002).
 31. J.P.P. Stewart, "MOPAC 7.0, A General Molecular Orbital Package", QCPE 455 (1993).
 32. W.L. Hase, R.J. Duchovic, X. Hu, K. Lim, D.-H. Lu, G.H. Peslherbe, K.N. Swamy, S.R. Vande Linde, H. Wang, R.J. Wolf, QCPE. VENUS is an enhanced version of MERCURY, W.L. Hase, QCPE 3, 453 (1983).
 33. G.H. Peslherbe, W.L. Hase, *J. Chem. Phys.* 104, 7882 (1996).
 34. J. Torras, E. Deumens, S.B. Trickey, *Journal of Computer-Aided Materials Design*, 13, 201, (2006).
 35. W. Smith, T.J. Forester, *J. Molec. Graphics*, 14, 136 (1996).
 36. W. Thiel, Program MNDO97 version 5.0.
 37. Object Management Group, OMG Common Object Request Broker Architecture: Core Specification. <http://www.omg.org/>
 38. C.E. Taylor, K. Runge, R.J. Bartlett, *Mol. Phys.*, 103, 2019 (2005).
 39. J.B. Foresman, C. Frisch, *Exploring Chemistry with Electronic Structure Methods*, 2nd ed., Gaussian Inc., Pittsburgh, PA. (1996).
 40. T. Steiner, A. M. M. Schreurs, J. A. Kanters, J. Kroon, *Acta Crystallogr. Sect. D*, 54, 25 (1998).
 41. J. L. Atwood, F. Hamada, K. D. Robinson, G.W. Orr, R. L. Vincent, *Nature*, 349, 683 (1991).
 42. G. Xiao, S. Liu, X. Ji, W. W. Johnson, J. Chen, J. F. Parsons, W. J. Stevens, G. L. Gilliland, R. N. Armstrong, *Biochemistry*, 35, 4753 (1996).
 43. P.W. Atkins and R.S. Friedman, «Molecular Quantum Mechanics»-3rd ed., Oxford Univ. Press, New York. (1997).
 44. Szabo, N, S. Ostlund, *Modern Quantum Chemistry*, Macmillan Publishing CO., INC, New York (1982).
 45. Κ. Α. Τσιπης Εισαγωγή στην Κβαντική Χημεία, Σημειώσεις για το Ελληνικό Ανοικτό Πανεπιστήμιο, Θεσσαλονίκη (2000).
 46. M. Born and J.R. Oppenheimer, *Ann. Physik*, 84, 457 (1927).
 47. M. N. Rosenbluth, A. W. Rosenbluth, *J. Chem. Phys.* 22, 881 (1954).
 48. C.C.J. Roothaan, *Rev. Mod Phys.*, 23, 69 (1951).
 49. P. W. Atkins and R. S. Friedman, *Molecular Quantum Mechanics*, Third edition, Oxford University Press, (1997).
 50. I. N. Levine, *Quantum Chemistry*, Fourth edition, Prentice-Hall, Englewood Cliffs, New Jersey (1991).



51. T. Pluta, A. J. Sadlej, *Chem. Phys. Let.*, 297,391 (1998).
52. J. C. Slater, *Phys. Rev.*, 36, 57 (1930).
53. S.F. Boys, *Pwc. R. Soc. (London)*, A200, 542 (1950).
54. Shavitt, *Modern theoretical chemistry*, 3, ed. H.F. Schaefer III, Plenum Press, New York, 189 (1977).
55. K. A. Brueckner, *Phys. Rev.*, 91, 1353 (1955). *Phys. Rev.*, 100, 36 (1955). H. P. Kelly, *Advan. Chem. Phys.*, 14, 129 (1969).
56. F. Coester, *Nucl. Phys.*, 1, 421 (1958). F. Coester and H. Kummel, *Nucl. Phys.*, 17, 477 (1960).
57. C. Meller and M.S. Plesset, *Phys. Rev.*, 46, 618 (1934).
58. B. Roos, *Lecture Notes of Theoretical Chemistry, Summer School, Oxford* (1974).
59. J. A. Pople, M. Head-Gordon and K. Raghavachari, *J. Chem. Phys.*, 87, 5968 (1987).
60. J. A. Pople, M. Head-Gordon and K. Raghavachari, *J. Chem. Phys.*, 90, 4635 (1989).
61. K. B. Lipkowitz and D. B. Boyd, *Reviews in Computational Chemistry III*, VCH publishers, New York (1992).
62. J. A. Nedler and R. Mead., *J. Computer*, 7, 308 (1965).
63. W. C. Davidon, *AEC Res. And Dev. Report. ANL-5990*.
64. R. Fletcher and M. J. D. Powell, *J. Computer*, 6, 163 (1963).
65. C. G. Broyden, *J. Inst. Maths. Applns.*, 6, 76 (1970).
66. R. Fletcher, *J. Computer*, 13, 317 (1970).
67. D. Goldfarb, *Maths. Comp.*, 24, 23 (1970).
68. D. F. Shanno, *Maths. Comp.*, 24, 647 (1970).
69. J.-M. Lehn, *Supramolecular Chemistry Concepts and Perspectives* (VCH, Weinheim, Germany, 1995).
70. M. Raimondi, G. Calderoni, A. Famulari, L. Raimondi, F. Cozzi, *J. Phys. Chem. A* 107, 772 (2003).
71. D.G. Papageorgiou, I.N. Demetropoulos, I.E. Lagaris, *Comput. Phys. Commun.* 109, 227 (1998).
72. D.G. Papageorgiou, I.N. Demetropoulos, I.E. Lagaris, *Comput. Phys. Commun.* 159, 70 (2004).
73. D.G. Papageorgiou, I.N. Demetropoulos, I.E. Lagaris, *Comput. Phys. Commun.* 109, 250 (1998).
74. F. G. Kalatzis, D. G. Papageorgiou, I. N. Demetropoulos, *Compur. Phys. Commun.* 175, 359 (2006). http://cpc.cs.qub.ac.uk/summaries/ADYB_v1_0.html
75. S.F. Boys and F. Bernardi, *Mol. Phys.* 19, 553 (1970).
76. R. M. Olson, M. W. Schmidt, M. S. Gordon, and A. P. Rendell. Enabling the efficient use of SMP clusters: The GAMESS/DDI model. *Proceedings of the 2003 ACM/IEEE Conference on Supercomputing*, p.41, November 15-21, (2003).
77. M.J.D. Powell, *TOLMIN: A Fortran package for linearly constrained optimization calculations*, Report No. DAMTP 1989/NA2, Univ. of Cambridge (1989).
78. M.J.D. Powell, *Math. Prog.* 45, 547 (1989).
79. D.G. Papageorgiou, I.N. Demetropoulos, I.E. Lagaris, *Melrin User Manual, CPC Program Library, Caralogue Identifier ADHQ*.
80. J. Stone, *The Theory of Intermolecular Forces* (Clarendon Press, Oxford, 1996).



81. R.J. Wheatley, A.S. Tulegenov, E. Bichoutskaia, *Int. Reviews in Physical Chemistry*, 23, 151 (2004).
82. London F, *Z. Phys.*, 63, 245 (1930).
83. London F, *Trans. Faraday Soc.* 33,8 (1937).
84. J.M. Jordy , A.D. Becke, *Chem. Phys. Lett.* 233, 134 (1995).
85. E.R. Johnson, G.A. DiLabio, *Chem. Phys. Lett.* 419, 333 (2006).
86. B. Jeziorski, R. Moszynski, K. Szalewicz, *Chem. Rev.* 94, 1887, (1994).
87. B. Jeziorski, R. Moszynski, A. Ratkiewicz, S. Rybak, K. Szalewicz, H.L. Williams, *SAPT: A Program for Many-Body Symmetry- Adapted Perturbation Theory Calculations of Intermolecular Interaction Energies*. In *Methods and Techniques in Computational Chemistry: METECC-94*; Clementi, E., Ed.; STEF: Gagliari, 1993; Vol. B.
88. B. Jeziorski, K. Szalewicz, and G. Chalasinski, *Int. J. Quantum Chem.* 14, 271 (1978).
89. G. Chalasinski and B. Jeziorski, *Mol. Phys.* 32, 81 (1976).
90. G. Chalasinski, B. Jeziorski, J. Andzelm, K. Szalewicz, *Mol. Phys.* 33, 971 (1977).
91. W. H. Adams, *Int. J. Quantum Chem.* 60, 273, (1996) and references therein.
92. E.M. Mas, K. Szalewicz, R. Bukowski, B. Jeziorski, *J. Chem.Phys.* 107, 4207 (1997).
93. F. van Duijneveldt, van Duijneveldt-van de Rijdt, G. C. J. M.; van Lenthe, J. H. *Chem. Rev.* 94. 1873, (1994).
94. G. Chalasinski, M.M. Szczesniak, *Chem. Rev.* 100, 4227, (2000).
95. B. Paizs, P. Salvador, A.G. Császár, M. Duran, S. Suhai, *J. Comput. Chem.* 22, 196 (2001).
96. *Mayer International Journal of Quantum Chemistry*, 100, 559 (2004).
97. Liu, B.; McLean, A. D. *J. Chem. Phys.* 59, 4557 (1973).
98. H. B. Jansen, P. Ross, *Chem Phys Lett* 3, 140 (1969).
99. I. Mayer, *Int J Quantum Chem.* 70, 41 (1998).
100. E. Gianinetti, M. Raimondi, E Tornaghi, *Int. J. Quant. Chem.* 60, 157 (1996).
101. Famulari, E. Gianinetti, M. Raimondi, M. Sironi, I Vandoni, *Theor. Chem. Acc.* 99, 358 (1998).
102. T. Nagata, O. Takahashi, K. Saito, S. Iwata, *J. Chem. Phys.* 115, 3553 (2001).
103. L. Mao, Y. Wang, Y. Liu, X. Hu, *J. Mol. Biol.* 336, 787 (2004).
104. S. Simon, M. Duran, J.J. Dannenberg, *J. Chem. Phys.* 105, 11024 (1996).
105. S. Suzuki, P.G. Green, R.E. Bumgarner, S. Dasgupta, W.A. Goddard III, G.A. Blake, *Science* 257, 942 (1992).
106. S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, *J. Am. Chem. Soc.* 122, 11450 (2000).
107. P. Hobza and R. Zahradnik, *Chem. Rev.* 88, 871 (1988).
108. E. R. Davidson and S. J. Chakravorty, *Chem. Phys. Lett.* 241,146 (1995).
109. J. W. Bevan, Z. Kisiel, A. C. Legon, D. J. Millen, and S. C. Rogers, *Proc. R. Soc. London, Ser. A* 372, 441 (1980).
110. S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, K. Tanabe, *J. Am. Chem. Soc.* 122, 11450 (2000).
111. S. Scheiner, T. Kar, J. Pattanayak, *J. Am. Chem. Soc.* 124, 13257 (2002).
112. R.N. Pribble, A.W. Garrett, K. Haber, T.S. Zwier, *J. Chem. Phys.* 103, 531 (1995).
113. G. A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, Oxford (1997).



114. J. Stone, *The Theory of Intermolecular Forces*, Clarendon, Oxford (1996).
115. B.O. Roos, P.R. Taylor, P.E.M. Siegbahn, *Chem. Phys.* 48, 157 (1980).
116. K. Ruedenberg, M.W. Schmidt, M.M. Gilbert, S.T. Elbert, *Chem. Phys.* 71, 41 (1982).
117. Szabados, Z. Rolik, G. Toth, P. R. Surjan, *J. Chem. Phys.* 122, 114104 (2005).
118. K. Andersson, P.-A. Malmqvist, B. O. Roos. *J. Chem. Phys.* 96, 1218 (1992).
119. K. Andersson, P.-A. Malmqvist, B. O. Roos, A. J. Sadlej, K. Wolinski, *J. Phys. Chem.* 94, 5483 (1990).
120. C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, J. P. Malrieu. *J. Chem. Phys.* 114, 10252 (2001).
121. C. Angeli, R. Cimiraglia, J. P. Malrieu. *Chem. Phys. Lett.* 350, 297 (2001).
122. C. Angeli, R. Cimiraglia, J. P. Malrieu. *J. Chem. Phys.* 117, 9138 (2002).
123. K. G. Dyall. *J. Chem. Phys.* 102, 4909 (1995).
124. Dalton, a molecular electronic structure program, Release 2.0 (2005), see <http://www.kjemi.uio.no/software/Dalton/Dalton.html>.
125. M. J. Watkins, K. Muller-Dethlefs, M. C. R. Cockett, *Phys. Chem. Chem. Phys.* 2, 5528 (2000).
126. G. Olaso-González, D. Roca-Sanjuán, L. Serrano-Andrés, M. Merchán, *J. Chem. Phys.* 125, 231102 (2006).
127. S. Suzuki, P.G. Green, R.E. Bumgarner, S. Dasgupta, W.A. Goddard III, G.A. Blake, *Science* 257, 942 (1992).
128. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, McGraw-Hill, 1982; R. McWeeny, *Methods of Molecular Quantum Mechanics*, Academic Press, 1992; W. J. Hehre, L. Radom, J. A. Pople and P. v. R. Schleyer, *Ab Initio Molecular Orbital Theory*, Wiley, 1986; J. Simons, *J. Phys. Chem.*, 95 1991, 1017; R. J. Bartlett, J. F. Stanton, *Rev. Comput. Chem.*, 5, 65 (1994).
129. Clementi E (ed) 1993 *Methods and Techniques in Computational Chemistry*, METECC-94 (Cagliari: STEF)
130. P.O. Lowdin *Adv. Chem. Phys.* 2, 207 (1959).
131. O. Sinanoglu, *Adv. Chem. Phys.* 6 315 (1964).
132. G.C. Lie, E. Clementi, *J. Chem. Phys.* 60, 1275 (1974).
133. G.C. Lie, E. Clementi, *J. Chem. Phys.* 60, 1288 (1974).
134. I. Krylov, C. D. Sherrill, E. F. C. Byrd, M. Head-Gordon *J. Chem. Phys.*, 109, 10669 (1998).
135. J. Frenkel, *Wave Mechanics, Advanced General Theory*, pp. 460–62. Oxford: Oxford Univ. Press (1934)
136. D.R. Hartree, W. Hartree, B. Swirles, *Philos. Trans. R. Soc. London Ser. A* 238:229 (1939).
137. T.L. Gilbert, *J. Chem. Phys.* 43, 248, (1965).
138. G. Das, A.C. Wahl, *J. Chem. Phys.* 44, 87 (1966).
139. R. Pauncz, *The Symmetric Group in Quantum Chemistry*. Boca Raton, FL: CRC, (1995).
140. E. Dalgaard, P. Jørgensen, *J. Chem. Phys.*, 69, 3833 (1978).
141. B. Datta B, D. Mukherjee, *Chem. Phys. Lett.*, 235, 31 (1995).
142. V. Glushkov, *Chem. Phys. Lett.*, 244, 1 (1995).
143. A.I. Krylov, C.D. Sherrill, E.F. Byrd, M. Head-Gordon, *J. Chem. Phys.*, 109, 10669 (1998).
144. X. Li, J. Paldus, *J. Chem. Phys.*, 110, 2844 (1999).
145. K. Nakayama, K. Hirao, R. Lindh, *Chem. Phys. Lett.*, 300, 303 (1999).
146. W.T. Borden, E.R. Davidson, *Acc. Chem. Res.*, 29, 67 (1996).



147. P. O. Lowdin, *Phys. Rev.*, 97, 1474 (1955).
148. D.M. Rogers, C. Wells, M. Joseph, V.J. Boddington, J.J.W. McDouall, *Journal of Molecular Structure*, 434, 239 (1998).
149. Brett Bode's MacMolPlt, available from
<http://www.msg.ameslab.gov/GAMESS/GAMESS.html>
150. W.T. Borden, E.R. Davidson, *Acc. Chem. Res.*, 29, 67 (1996).
151. K. Wolinski, H.L. Sellers, P. Pulay, *Chem. Phys. Lett.* 140, 225 (1987).
152. H.J. Werner *Mol. Phys.* 89, 645 (1996).
153. R.B. Murphy, R.P. Messmer, *Chem. Phys. Lett.* 183, 443 (1991).
154. M. R. Hoffmann, *Chem. Phys. Lett.* 210, 193 (1993).
155. H. Nakano, *Chem. Phys. Lett.* 207, 372 (1993).
156. K. Anderson, P.-Å. Malmqvist, B.O. Roos, A.J. Sadlej and K. Wolinski: *J. Phys. Chem.* 94, 5483 (1990).
157. K. Anderson, P.-Å. Malmqvist and B.O. Roos: *J. Chem. Phys.* 96, 1218 (1992).
158. K. Hirao, *Chem. Phys. Lett.* 190, 374 (1992).
159. K. Hirao, *Chem. Phys. Lett.* 196, 397 (1992).
160. H. Nakano, K. Nakayama, K. Hirao, M. Dupuis, *J. Chem. Phys.* 106, 4912 (1997).
161. T. Hashimoto, H. Nakano, K. Hirao, *J. Mol. Struct. (THEOCHEM)* 451, 25 (1998).
162. B.O. Roos, M. Fülcher, P.A. Malmqvist, M. Merchán and L. Serrano-Andrés: *Theoretical Studies of the Electronic Spectra of Organic Molecules*, in S.R. Langhoff (ed.) *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, Kluwer Academic Publishers, Dordrecht, (1995).
163. R. B. Murphy and R. P. Messmer, *Chem. Phys. Letters* 183, 443 (1991).
164. H.-J. Werner, *Mol. Phys.* 89, 645 (1996).
165. M.S. Gordon, M.W. Schmidt, G.M. Chaban, K.R. Glaesemann, W.J. Stevens, C. Gonzalez, *J. Chem. Phys.*, 110, 4199 (1999).
166. K. I. Peterson, W. Klemperer. *J. Chem. Phys.* 85, 725 (1986).
167. B. Kovacevic, D. Baric, Z.B. Maksic, T. Muller, *ChemPhysChem* 5, 1352 (2004).
168. Engdahl and B. Nelander *Chem. Phys. Lett.* 113, 49 (1985).
169. P.O. Widmark, P.A. Malmqvist, B. Roos, *Theor. Chim. Acta* 77, 291 (1990).
170. K. B. Wiberg, *J. Comp. Chem.* 25, 1342 (2004).
171. P. Tarakeshwar, H. S. Choi, S. J. Lee, J. Y. Lee, K. S. Kim, T. K. Ha, J. H. Jang, J. G. Lee, H. Lee, *J. Chem. Phys.* 111, 5838 (1999).
172. Courty, M. Mons, I. Dimicoli, F. Piuze, M.-P. Gaigeot, V. Brenner, P. Pujo, P. Millie, *J. Phys. Chem. A* 102, 6590 (1998).
173. T. Otsuka, H. Nakai, *J. Comput. Chem.* 28, 1137 (2007).
174. M.H. Kim, L. Shen, H. Tao, T.J. Martinez, A.G. Suits, *Science*, 315, 1561 (2007).
175. M. P. Allen and D.J. Tildesley, «*Computer Simulation of Liquids*», Oxford Univ. Press, New York. (1992).
176. W. F. van Gunsteren and H.J.C. Berendsen, *Angew. Chem. Int. Ed Engl.*, 29, 992 (1990).
177. P. Ren and J. W. Ponder, *J. Phys. Chem. B*, 107, 5933-5947 (2003). P. Ren and J. W. Ponder, *J. Comput. Chem.*, 23, 1497-1506 (2002). R. V. Pappu, R. K. Hart and J. W. Ponder, *J. Phys. Chem. B*, 102, 9725-9742 (1998). M. E. Hodsdon, J. W. Ponder and D. P. Cistola, *J. Mol. Biol.*, 264, 585-602 (1996). C. E. Kundrot, J. W. Ponder and F. M. Richards, *J. Comput. Chem.*, 12, 402-409 (1991). J. W. Ponder and F. M. Richards, *J. Comput. Chem.*, 8, 1016-1024 (1987).
178. <http://dasher.wustl.edu/tinker/summary.html>



179. B. R. Brooks, "Algorithms for Molecular Dynamics at Constant Temperature and Pressure", DCRT Report, NIH, April 1988.
180. L. Verlet, Phys. Rev., 159, 98 (1967).
181. L. Verlet, Phys. Rev., 165, 201 (1968).
182. H. J. C. Berendsen, J.P.M Postma, W.F.v. Gunsteren, A. DiNola and J.R. Haak, J. Chem. Phys., 81, 3684 (1984).
183. Physics and Astronomy Classification Scheme (pacs); see <http://publish.aps.org/PACS>
184. R. Car and M. Parrinello, Phys. Rev. Lett., 55, 2471 (1985).
185. B. H. Schlegel, Bull. Korean Chem. Soc., 24, 6, (2003).
186. Helgaker, T.; Uggerud, E.; Jensen, H. J. A. Chem. Phys. Lett., 173, 145 (1990).
187. Uggerud, E.; Helgaker, T. J. Am. Chem. Soc., 114, 4265 (1992).
188. K. Bolton, W. L. Hase, and G. H. Peslherbe, in Modern Methods for Multidimensional Dynamics Computations in Chemistry, ed. D. L. Thompson (World Scientific, Singapore, 1998).
189. G. H. Peslherbe, W. L. Hase, J. Chem. Phys., 104, 7882 (1996).
190. C. L. Brooks, D. Case, Chem. Rev. 93, 2487 (1993).
191. T. J. Marrone, M. K. Gilson, J. A. McCammon, J. Phys. Chem. 100, 1439 (1996).
192. D. Wei, H. Guo, D. R. Salahub, Phys. Review E, 64, 11907 (2001).
193. J. P. Perdew, Phys. Rev. B 33, 8822 (1986). A. D. Becke, Phys. Rev. A 38, 3098 (1988).
194. St-Amant, D. R. Salahub, Chem. Phys. Lett. 169, 387 (1990).
195. Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. "Universal Reaction Field Model Based on Ab Initio Hartree-Fock Theory" Chem. Phys. Lett, 288, 293 (1998).
196. C. Chamberlin, J. Pu, C. P. Kelly, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.-Y. Chuang, P. L. Fast, B. J. Lynch, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, D. G. Truhlar, GAMESSPLUS-version 4.7, University of Minnesota, Minneapolis, 2005, based on the General Atomic and Molecular Electronic Structure System (GAMESS) as described in M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, J. Comput. Chem. 14, 1347 (1993). <http://comp.chem.umn.edu/gamesplus/>
197. Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. "A New Class IV Charge Model for Extracting Accurate Partial Charges from Wave Functions" J. Phys. Chem. A, 102, 1820 (1998).
198. D.B. McGarrah, R.S. Judson, Journal of Computational Chemistry, 14, 11, 1385 (1993).
199. H. Kessler, Angew. Chem. Int. Ed. Engl., 21, 512 (1982).
200. D.G.H Ballard, C.H. Bamford, F.J. Weymouth, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 227, 1169, page 155, (1955).
201. V. A. Tasis, A. Stavrakoudis, I. N. Demetropoulos, Molec. Simulations, 32, 643 (2006).

