

DEMOCRITUS UNIVERSITY OF THRACE

SCHOOL OF HEALTH SCIENCES

DEPARTMENT OF MOLECULAR BIOLOGY & GENETICS

NOEmd :

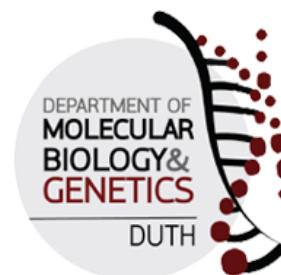
A computer program to automate the comparison between experimental NMR data and molecular dynamics simulations

BSc Thesis

Supervisor:

Nicholas M. Glykos
Professor of Structural And Computational Biology
Dept. of Molecular Biology and Genetics
Democritus University of Thrace

Tsevis Spyridon-Aggelos, Registration Number: 2524



DEMOCRITUS UNIVERSITY OF THRACE

SCHOOL OF HEALTH SCIENCES

DEPARTMENT OF MOLECULAR BIOLOGY & GENETICS

NOEmd : A computer program to automate the comparison between experimental NMR data and molecular dynamics simulations

BSc Thesis

I declare that the present thesis entitled ‘NOEmd : A computer program to automate the comparison between experimental NMR data and molecular dynamics simulations’ is original and was carried out by me personally, as an undergraduate student of the Department of Molecular Biology and Genetics, with Registration Number 2524. I certify that during the preparation and writing of the thesis, all legal requirements were followed, and that the principles of academic ethics and integrity were fully adhered to, which prohibit the falsification of results, the misuse of others' intellectual property, and plagiarism.

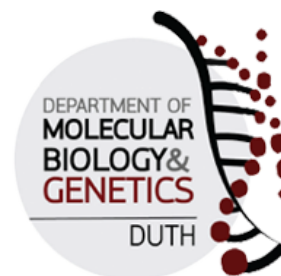
Tsevis Spyridon-Aggelos, Registration Number: 2524

Supervisor:

Nicholas M. Glykos

Associate Professor, Laboratory of Structural and Computational Biology

Democritus University of Thrace



ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ

ΤΜΗΜΑ ΜΟΡΙΑΚΗΣ ΒΙΟΛΟΓΙΑΣ & ΓΕΝΕΤΙΚΗΣ

NOEmd : Ένα υπολογιστικό πρόγραμμα για την αυτοματοποίηση της σύγκρισης δεδομένων NMR και προσομοιώσεων μοριακών δυναμικών

BSc Thesis

Δηλώνω ότι η παρούσα εργασία με τίτλο 'NOEmd : Ένα υπολογιστικό πρόγραμμα για την αυτοματοποίηση της σύγκρισης δεδομένων NMR και προσομοιώσεων μοριακών δυναμικών' είναι πρωτότυπη και πραγματοποιήθηκε από εμένα προσωπικά, προπτυχιακό φοιτητή του Τμήματος Μοριακής Βιολογίας και Γενετικής, με Αριθμό Μητρώου 2524. Βεβαιώνω ότι κατά την εκπόνηση της εργασίας και τη συγγραφή της τηρήθηκαν τα προβλεπόμενα από το νόμο, καθώς και ότι ακολουθήθηκαν πλήρως οι αρχές της ακαδημαϊκής ηθικής και δεοντολογίας, οι οποίες απαγορεύουν την παραποίηση των αποτελεσμάτων, την κατάχρηση της διανοητικής ιδιοκτησίας άλλων και τη λογοκλοπή.

Τσέβης Σπυρίδων-Άγγελος, Αριθμός Μητρώου: 2524

Επιβλέπων:

Νικόλαος Μ. Γλυκός

Αναπληρωτής Καθηγητής, Εργαστήριο Δομικής και Υπολογιστικής Βιολογίας, ΔΠΘ

Acknowledgements

This thesis would not have been possible without the guidance, emotional support, and encouragement of many of those around me. I know some will feel left out, but even those that go unmentioned know how much of a profound impact they have left on me.

First and foremost, I would be remiss if I didn't express my deepest thanks and gratitude to my supervisor, Professor N. M. Glykos. Your mentorship, assistance with any major or minor trouble, and continuous belief in my potential have been the foundational pillars of both my studies so far as well this particular thesis. Thank you for being the catalyst for every ounce of intellectual curiosity I've picked up on my journey as a student and for pushing me to become a better scientist. Your guidance through the complexities of research and struggles of development has been most valuable and I will be forever grateful to you for giving me a chance to work alongside you.

Secondly I am also incredibly grateful for my friends, who have always been a source of support and encouragement. They kept me sane through the late nights of work and everyday challenges, as well as everything else life brought my way, our friendship kept my spirits high every step of the way. Thank you for the late night exits, the laughter, and for always being there to listen.

Lastly, I owe everything as well as my greatest thanks to my family. Your unconditional love, care, and support shaped me to be the person I am today. Thank you for believing in me even when my own belief tank was running low on fuel, for celebrating my successes, and for providing comfort during the times when it was most needed. This achievement belongs to you just as much as it does to me.

Abstract

To understand the function of biomolecules, we must first determine their three-dimensional structures. Nuclear Magnetic Resonance (NMR) spectroscopy is a crucial technique for studying molecular structures in solution, and it was first observed in 1944 by Nobel Laureate Isidor Rabi. The nuclear overhauser effect (NOE) is used as a powerful structural anchor to determine inter-atomic distances, but the translation of NOE intensities into restraints and fitting them to molecular dynamics (MD) simulations is prone to error, specifically when structurally equivalent protons have degenerate NOEs.

The Python code of NOEmd [<https://github.com/tsevisaggelos/nOeMD/>], provided with this dissertation, presents a unique framework that simplifies the analysis of NOE-derived distance restraints, automating the process so that NOEs can be used with ease for biomolecular modeling, validation, refinement, and structure analysis. It uses multiple, commonly known and used file formats to parse the molecular data (.mr), the protein structure (.psf), and the simulation files (.dcd), producing an accurate report of the experimental and restraint data from the NOEs. As a result, this research provides a fast way to parse experimental restraint files using a mapping based on established nomenclature of the atom names to allow for a more reliable representation of the protein structure and a consistent mapping of the experimental restraints with the simulation data. In addition, the software offers the ability to deal with degenerate restraints in an accurate manner, as it uses chemical and structural information to define and categorize the different degenerate groups, using these groups to filter out the correct distances of the simulation for subsequent comparisons. Finally, it offers a thorough NOE-violation analysis as the program calculates both r^{-3} and r^{-6} ensemble averages from the simulation, comparing these against the experimental upper bound distances to produce a detailed and informative report listing the deviation and violation of each restraint. Using standard Python libraries such as NumPy, it attempts to efficiently process these averages, in a resource-lite manner. The program also uses the carma commands [Glykos, N.M. (2006), Koukos, P.I. & Glykos, N.M. (2013)] to read the MD files and produce the distances required for calculating the ensemble averages for the analysis. With the use of the NOEmd script, it is demonstrated that the manual effort required for NOE analysis is significantly reduced, and that this process becomes significantly less complicated and tedious, while also resulting in accurate findings.

Περίληψη

Για να κατανοήσουμε τη λειτουργία των βιομορίων, πρέπει πρώτα να προσδιορίσουμε τις τρισδιάστατες δομές τους. Η φασματοσκοπία πυρηνικού μαγνητικού συντονισμού (NMR) αποτελεί μια σημαντική τεχνική για τη μελέτη μοριακών δομών σε διάλυμα, και παρατηρήθηκε για πρώτη φορά το 1944 από τον κάτοχο του βραβείου Νόμπελ Isidor Rabi. Το φαινόμενο Nuclear Overhauser (NOE) χρησιμοποιείται ως ισχυρό δομικό σημείο αναφοράς για τον προσδιορισμό των διατομικών αποστάσεων, αλλά η μετατροπή των εντάσεων του NOE σε restraints και η προσαρμογή τους σε προσομοιώσεις μοριακών δυναμικών (MD) είναι επιρρεπής σε σφάλματα, ειδικά όταν δομικά ισοδύναμα πρωτόνια παρουσιάζουν εκφυλισμένα NOE.

Ο κώδικας Python του NOEmd [<https://github.com/tsevisaggelos/nOeMD/>], ο οποίος παρέχεται μαζί με την παρούσα πτυχιική, παρουσιάζει ένα μοναδικό πλαίσιο που απλοποιεί την ανάλυση των restraints απόστασης που προέρχονται από το NOE, αυτοματοποιώντας τη διαδικασία, έτσι ώστε τα NOE να μπορούν να χρησιμοποιηθούν με ευκολία για τη μοντελοποίηση, την επικύρωση, τη βελτίωση και την ανάλυση της δομής βιομορίων. Χρησιμοποιεί πολλαπλές, ευρέως γνωστές μορφές αρχείων για την ανάλυση των μοριακών δεδομένων (.mr), της δομής της πρωτεΐνης (.psf) και των αρχείων προσομοίωσης (.dcd), παράγοντας μια ακριβή αναφορά των πειραματικών δεδομένων και των δεδομένων περιορισμών που προέρχονται από τα NOE. Ως αποτέλεσμα, η παρούσα έρευνα παρέχει έναν γρήγορο τρόπο ανάλυσης των αρχείων πειραματικών περιορισμών χρησιμοποιώντας μια αντιστοίχιση των ονομάτων των ατόμων βασισμένη στην ονοματολογία που έχει εδραιωθεί, ώστε να επιτρέπεται μια πιο αξιόπιστη αναπαράσταση της δομής της πρωτεΐνης και μια συνεπής αντιστοίχιση των πειραματικών περιορισμών με τα δεδομένα προσομοίωσης. Επιπλέον, το λογισμικό προσφέρει τη δυνατότητα ακριβούς χειρισμού των εκφυλισμένων restraints, καθώς χρησιμοποιεί χημικές και δομικές πληροφορίες για τον ορισμό και την κατηγοριοποίηση των διαφόρων εκφυλισμένων ομάδων, αξιοποιώντας αυτές τις ομάδες για να φιλτράρει τις σωστές αποστάσεις της προσομοίωσης με σκοπό τις επακόλουθες συγκρίσεις. Τέλος, προσφέρει μια διεξοδική ανάλυση των restraints του κανόνα NOE, καθώς το πρόγραμμα υπολογίζει τόσο τους μέσους όρους συνόλου r^{-3} όσο και r^{-6} από την προσομοίωση, συγκρίνοντάς τους με τις πειραματικές ανώτατες αποστάσεις, προκειμένου να παράγει μια λεπτομερή και κατατοπιστική αναφορά που απαριθμεί την απόκλιση και την παραβίαση κάθε restraint. Χρησιμοποιώντας τυπικές βιβλιοθήκες Python, όπως το NumPy, επιχειρεί να επεξεργαστεί αποτελεσματικά αυτούς τους μέσους όρους, με τρόπο που απαιτεί ελάχιστους πόρους. Το πρόγραμμα χρησιμοποιεί επίσης εντολές της carma [Glykos, N.M. (2006), Koukos, P.I. & Glykos, N.M. (2013)] για την ανάγνωση των αρχείων MD και την παραγωγή των αποστάσεων που απαιτούνται για τον υπολογισμό των μέσων όρων του συνόλου για την ανάλυση. Με τη χρήση του script του NOEmd, αποδεικνύεται ότι η χειροκίνητη εργασία που απαιτείται για την ανάλυση NOE μειώνεται σημαντικά και ότι η διαδικασία αυτή γίνεται πολύ λιγότερο περίπλοκη και κουραστική, ενώ παράλληλα οδηγεί σε ανάλογα ακριβή αποτελέσματα.

Table of Contents

Abstract.....	5
Περίληψη.....	6
Chapter 1: Introduction.....	9
1.1 The Role of Structural Biology.....	9
1.2 NMR Spectroscopy for Structure Determination.....	9
1.3 Challenges in NOE Data Analysis.....	9
1.4 Thesis Overview: The NOEmd Framework.....	10
Chapter 2 - Methods: Parsing and Mapping of Structural Data.....	11
2.1 Introduction.....	11
2.2 Parsing the Experimental Restraint File (.mr).....	11
Code Excerpt 2.1: parse_mr_file.....	11
2.3 Parsing the Protein Structure File (.psf).....	12
2.4 The Markley Nomenclature and Degeneracy Mapping.....	12
Code Excerpt 2.2: The markley_map and reverse_markley_map.....	12
2.5 Generating the Master Atom Pair Table.....	14
Chapter 3 - Methods: Interatomic Distance Calculation & Degeneracy Resolution 14	14
3.1 Introduction.....	14
3.2 Interfacing with carma for Distance Calculation.....	14
3.3 The Challenge of Degenerate Restraints.....	15
3.4 The calculate_noe_sums_in_memory Function.....	15
Code Excerpt 3.1: calculate_noe_sums_in_memory.....	16
Chapter 4 - Methods: NOE Averaging and Violation Analysis.....	18
4.1 Introduction.....	18
4.2 The Theory of NOE Averaging.....	18
4.3 The calculate_noe_sums_in_memory Function.....	18
Code Excerpt 4.1: Running Sum Accumulation in calculate_noe_sums_in_memory.....	18
4.4 The run_noe_averaging Function: Generating the Final Report.....	20
Chapter 5 - Results: Case Study - Analysis of the Chignolin Variant CLN025.....	22
Comparative Result Analysis: CLN025 Trajectories.....	31
Summary Statistics.....	31
Key Structural Differences and Variations.....	31
1. The Core Divergent Anchor: Tyr2 – Pro4.....	31
2. Long-Range Aromatic Packing: Tyr2 – Trp9.....	32
3. Backbone Terminal Regions.....	32
Chapter 6 - Results: Case Study 2 - Analysis of 2N0X.....	36
Comparison of Representative MD Clusters with the Experimental NMR Structure.....	37
Core Divergence Insights.....	39
1. The Ensemble Averaging Effect (r-6 Weighting).....	39
2. Extreme Outliers in Specific Clusters (DPCA02 & DPCA06).....	39
3. Population and Resolution Trends.....	40

Chapter 7: Conclusion and Final Remarks.....	64
7.1 Summary of Contributions.....	64
7.2 Final Remarks.....	65
Source Code.....	66
References.....	94

Chapter 1: Introduction

1.1 The Role of Structural Biology

- Biological macromolecules are functional to the extent that they can be accurately folded in the three-dimensional space. The atomic structure of a protein or nucleic acid determines how it binds to its ligands, how it acts as an enzyme, how it transduces signals, and how it works with other molecular species in biological pathways. Determining these structures, ideally at single-atom resolution, is thus critical for a broad range of studies in biochemistry and molecular biology. X-ray crystallography, cryo-electron microscopy (cryo-EM), and nuclear magnetic resonance (NMR) spectroscopy have enabled determination of most protein structures. Testing theoretical coordinate ensembles against experimental observables, including NMR NOE distance restraints, is a critical element of bioinformatics studies in the fields of structural biology and biophysics. While Molecular Dynamics (MD) trajectories have much better time resolution than current NMR structure determination methods, the accuracy of such simulations is limited both by the empirical nature of force fields and the time-consuming nature of conformational sampling. Thus, it is necessary to test theoretical MD trajectories against experimental observables to determine the physical accuracy of computational approaches.

1.2 NMR Spectroscopy for Structure Determination

In contrast to X-ray crystallography and cryo-EM, NMR spectroscopy allows for the chance to study molecules in the solution state, much like they exist physiologically. This means one can also see how biomolecules behave dynamically, so you can get an idea of their flexibility, conformational changes and intermolecular interactions. A main source of structural information in NMR is NOE (Nuclear Overhauser Effect), which involves the transfer of nuclear spin polarization from one nucleus to another across space. As NOE effects are highly distance-dependent, and approximately proportional to r^{-6} , NMR measures short range interatomic distances (typically less than 6 Å). By collecting a large data set of proton-proton distances restraints the structure can be calculated (Neuhaus & Williamson 2000).

1.3 Challenges in NOE Data Analysis

Transforming raw NOE spectral data into a good structural model typically consists of several steps, which can often be fraught with difficulty.

1. **Assignment and Nomenclature:** As the first step, the NOE cross-peaks are assigned to the proton pairs in the molecule. This assignment process can be a daunting task because of the overlapping peaks and ambiguities. Also, there is variation in naming of atoms due to different conventions used by the different softwares and labs. For standardizing the name of each different proton, especially those found in the methyl and methylene groups, the Markley nomenclature is used (Markley et al., 1998).
2. **Degeneracy:** Many of the protons in a protein are chemically and structurally equivalent. The three protons in a methyl group, for instance, cannot be discriminated from one another in most NMR experiments. An NOE to a methyl group could be due to any of its three protons. We therefore refer to the three protons as degenerate. A correct treatment of the NOE must take this degeneracy into account. The NOE cannot be written as a restraint to one of the three protons, since this would not be correct. The effective distance is the population-weighted distance to all the equivalent protons.
3. Since NOE values represent an average over time, they need to be calculated as the ensemble average over all of the different possible configurations of the molecule. This is a consequence of the fact that biomolecules in solution are constantly in motion; this movement includes thermal motions which occur on the order of the NMR timescale. As a result, it is difficult to validate a given structural model in an MD simulation by using the NOEs that would have been observed during an experiment (Case, 1999).
4. **Workflow Integration:** Our analysis workflow uses multiple distinct data files, the experimental restraints .mr file, system topology .psf file, and simulation trajectory .dcd file, all from different sources. Bringing them together and doing the appropriate analysis by hand is complicated, sometimes tedious, and can lead to mistakes.

1.4 Thesis Overview: The NOEmd Framework

To overcome the challenges listed above, this paper presents a new NOE-restraint-analysis program called NOEmd that automates the entire analysis procedure. It functions as a complete script with standardized input files that processes data from scratch, all the way to an overall quantitative violation summary. These next few chapters describe the overall architecture of NOEmd in depth:

- The **second chapter** will explain the procedures of parsing and mapping structural data and how the script reads the .mr and .psf files to disambiguate the atom names through documented nomenclature.
- **Chapter 3** outlines calculating pairwise interatomic distances with subprocess interfacing with carma, and how the script dynamically resolves proton degeneracies

at runtime by using volatile system RAM to load in the necessary data columns without the overhead of reading through intermediary files.

- In the **fourth chapter**, we delve into the analytical part of the program, describing the progressive update of the r^3 and r^6 running sums during a single pass through a file, the calculation of ensemble time averages based on a specific formula, and the output of the comprehensive color-coded output analysis of the program.

Through this automation, NOEmd serves as a fast and dependable tool for more expedient and thorough validation of biomolecular structures versus experimental NMR data.

Chapter 2 - Methods: Parsing and Mapping of Structural Data

2.1 Introduction

Interpreting input data accurately and succinctly from different data sources constitutes the most basic aspect of any automated analysis software. Regarding NOE analysis, the main data inputs will usually be in the form of an experimental restraints file, providing the distance information, and the protein structure file, providing all the atomic content and connectivity of the structure. The NOEmd framework is compatible with the commonly used .mr and .psf file types. At the data interpretation stage, there is also the crucial need to define a naming convention that will be used to map the atom names between simulation and experimental data.

2.2 Parsing the Experimental Restraint File (.mr)

The .mr file constitutes a list of experimentally extracted distance restraints. Each restraint defines a pair of atoms and the upper bound for the distance between the two atoms. The function for reading this file, `parse_mr_file`, has been contracted in the source script.

Code Excerpt 2.1: `parse_mr_file`

```
def parse_mr_file(mr_file_path):
    """Returns atom pairs AND their experimental distances."""
    entries = []
    entry_pattern = re.compile(

r"(\d+)\s+(\w+)\s+(\w+)\s+(\d+)\s+(\w+)\s+(\w+)\s+([\d.]+)"
    )
    with open(mr_file_path, 'r') as mr_file:
```

```

    for line in mr_file:
        if line.startswith('*') or not line.strip():
            continue
        match = entry_pattern.search(line)
        if match:
            atom1 = (match.group(1), match.group(2),
match.group(3))
            atom2 = (match.group(4), match.group(5),
match.group(6))
            distance = float(match.group(7))
            entries.append((atom1, atom2, distance))
    return entries

```

Explanation: This function uses a pre-compiled regular expression to achieve the desired results. Lines are then iterated, and comments and blank lines are ignored. The compiled regular expression object is used to find a match of seven groups of characters: residue number and name of the first residue, residue number and name of the second residue, atom name of the first atom, atom name of the second atom, and the measured distance. When the regular expression does find a match, all characters belonging to each match are claimed and grouped into variables, and each of these are appended to a list of all the entries and this list is returned to the function that called it.

2.3 Parsing the Protein Structure File (.psf)

In CHARMM and NAMD, the .psf file contains an entry for each atom of the molecule. This file contains information on the ID, type, charge, and connectivity for each atom; it is thus a “master” file. The function `find_all_atom_entries` searches the .psf file looking for an ID of an atom from its residue number, atom name, and residue number. Atom naming might change, for example the protons on the methyl group can be designated as HB1, HB2, and HB3 in a .psf file while they should be QB in a .mr file.

2.4 The Markley Nomenclature and Degeneracy Mapping

All this ambiguity is resolved by the use of a mapping table (also known as a dictionary) based on the official Markley nomenclature defined by IUPAC-IUBMB. The Markley map, (markley_map), has been manually compiled with information on which pseudo-atom names represent sets of atoms: QG (a pseudo-atom for a degenerate methyl) would map to HG11, HG12, HG13, etc.

Code Excerpt 2.2: The `markley_map` and `reverse_markley_map`

```
# Amino acid specific mappings from Markley nomenclature
markley_map = {
    "ALA": {"QB": ["HB1", "HB2", "HB3"]},
    "VAL": {
        "QG1": ["HG11", "HG12", "HG13"],
        "QG2": ["HG21", "HG22", "HG23"]
    },
    # ... and so on for all relevant amino acids
}

# Generate reverse mappings
reverse_markley_map = {}
for res_name, mappings in markley_map.items():
    res_reverse_map = {}
    for k, v in mappings.items():
        if isinstance(v, list):
            for item in v:
                res_reverse_map[item] = k
        else:
            res_reverse_map[v] = k
    reverse_markley_map[res_name] = res_reverse_map
```

Explanation: The `markley_map` takes each pseudo-atom name, for example "QB" for Alanine, and maps it directly to a list of actual proton names. Subsequently, we create the `reverse_markley_map` so this mapping can work omnidirectionally. In case the input files use either pseudo-atom names or proton names, it is useful to have the safeguard that either way, the correct atom is identified. In the process of finding the entry of the atom, a search of the .psf file is conducted using the following set of possibilities, `atom_types_to_try`, which are derived from forward and backward maps, and so if a .mr file contains a restraint of atom VAL QG1, the .psf entry corresponding to atom HG11 is found. And the same applies for HG12 and HG13.

2.5 Generating the Master Atom Pair Table

The `find_atom_ids_for_pairs` function will go through the entries that are stored in the memory to find a match. To make the process more efficient it first collects all of the atoms it has encountered in the restrained lines and it will store those in a `unique_atoms` set that prevents the function from repeatedly searching within a file.

It then calls `find_all_atom_entries()` on every atom in the `unique_atoms` set, thereby storing a local in-memory cache called `atom_entries_cache`. The function iterates through all the entries of interest and stores that in `atom_entries_cache`. The function will also keep a `seen_pairs` set to track the found matched pairs to then sort them so that redundant cross assignments to unique IDs of atoms are prevented.

It stores the found atomic id mappings in `formatted_atom_pairs` set and checks if the set is empty before passing the atomic id set and error flag named `has_unknown` to the other steps.

Finally the printing of the master table to the screen and also the output file `table1.out` (also a printout of the first table) is made in the `main()` function at the end of the script. Any discrepancy, for example, an atom that does not exist in structure data will trigger a script abort in the form of an error and terminate the program in `sys.exit(1)` in order to prevent the expensive computation load from going ahead which could also possibly result in memory overflow at the end of the program.

Chapter 3 - Methods: Interatomic Distance Calculation and Degeneracy Resolution

3.1 Introduction

Now that we've identified the restraint atom IDs for each experimental restraint in our protein file, we'll calculate how well a set of structures, e.g. from an MD simulation, satisfies the set of restraints. In particular, for each frame of a simulation, we need to calculate the pairwise atom distances for each restraints. How NOEmd computes these distances and, very

importantly, how the calculated distances are postprocessed so that we can accurately and efficiently handle degenerate restraints in a memory-conservative manner, will be described below.

3.2 Interfacing with carma for Distance Calculation

As some may have surmised, evaluating pair distances for thousands of atoms over the course of thousands or more trajectory frames is a fairly costly and heavy-duty operation. Instead of building the required functionality from scratch, the script relies on the well optimized stand alone program carma, which was written by my supervisor N. M. Glykos for use in biomolecular simulation studies of trajectory analysis. The framework first reserves a single, stand alone text resource in the working directory called carmainput_for_distcalc.txt and uses that as a registry for the internal execution flow in carma. This file contains a sorted list of pairs of atom IDs that match a master restraint map file. The script calls carma as an external process and runs the commands one at a time. Carma is handed the .psf file which contains topology information, the .dcd file which contains the simulation trajectory, and the newly created file containing the list of pairs to be calculated. Since carma understands the coordinates within the trajectory frames, all it needs to do is dump the resulting distance data to a single output file called carma.distances. At the end of the calculation, the script then deletes the text file that was used as an execution registry, carmainput_for_distcalc.txt, in order to avoid unnecessary clutter in the working directory.

3.3 The Challenge of Degenerate Restraints

The carma.distances file stores an array matrix, in which each column refers to unique atom-pair restraints and each row refers to the successive simulation frames. This raw data trace is not useful for analyses directly, since the physical degeneracy of the proton has not been taken into account. For example: an NOE between a backbone proton and a valine QG1 methyl group has the corresponding translational mapping of the QG1 target pseudo-atom into 3 actual chemical proton atoms.

Consequently, carma will yield 3 distinct spatio-temporal distance arrays, all corresponding to the same experimental NOE restraints. The observed intensity of a local NOE is related to the ensemble time-average of r^{-6} , in which “r” is the distance between two protons, implying that the shorter distances have the most influence on the NOE.

Since the magnitude of a local NOE cross-peaks intensity scales with the ensemble time-averaged of (r^{-6}) , small fluctuations in “r” are more significant, in that their value of (r^{-6}) will be higher; in fact, as is often the case, the smallest (r) values will contribute the majority to the sum (Σ/r^6). Thus, a fail-proof way of distinguishing (per frame) equivalent pairs is the shortest (r) value among those pairs.

3.4 The `calculate_noe_sums_in_memory` Function

These multi-column degeneracies are avoided in the program design, and to further optimize the runtime of the distance matrices a space-efficient computational, an exclusive memory-dependent step has been implemented within the `calculate_noe_sums_in_memory` function. As a result, file IO operations and disk-based text parsing and filtering is eliminated, leading to faster processing speed and reduced runtime.

Code Excerpt 3.1: `calculate_noe_sums_in_memory`

```
def calculate_noe_sums_in_memory(formatted_atom_pairs,
    carma_file='carma.distances'):
    """
    Reads the raw carma.distances file, resolves degeneracies for
    each frame
    in memory, and directly calculates the sum of  $r^{-3}$  and  $r^{-6}$ 
    without
    writing an intermediate file.
    """
    # Step 1: Group columns by their restraint definition.
    column_groups = defaultdict(list)
    unique_restraint_keys = []
    raw_col_to_unique_group_idx = {}

    temp_unique_id_pairs_for_carma = set()
    col_idx_counter = 0
    for pair_desc in formatted_atom_pairs:
        # ... [Parses columns & isolates matching ID pairs] ...
        # Define the group key using residue/atom names
        k1 = (left_psf[2], left_psf[3], left_psf[4])
        k2 = (right_psf[2], right_psf[3], right_psf[4])
        group_key = tuple(sorted((k1, k2)))

        if group_key not in column_groups:
            unique_restraint_keys.append(group_key)

        group_idx = unique_restraint_keys.index(group_key)
        column_groups[group_key].append(col_idx_counter)
        raw_col_to_unique_group_idx[col_idx_counter] = group_idx
        col_idx_counter += 1
```

Explanation - Dynamic Restraint Grouping: It retrieves from context's internal `formatted_atom_pairs` list residue sequence numbers, residue names and original atom-types and links them together into a `group_key` set that maps all the individual degenerate atom-pair columns in the distance file to the original single experimental restraint.

- **Optimized File Parsing:** The script opens the `carma.distances` file and reads it line by line (one frame per line). It skips the frame number on each line and reads the values as the distances.
- **On-The-Fly Minimum Resolution:** Instead of caching a large text matrix for storing the computed result or writing intermediate files to disk, we simply process the frame by using only system memory. When running the loop over each compiled restraint group for the given frame, the user retrieves the distance entries that correspond to the current frame group from the raw array containing the frame data. Subsequently, the minimum distance is identified with the built-in function `min()` in the Python language.
- The minimum distance between each degenerate group in the frame is directly used to update the tracking variable matrices and is also included in a running numerical total. These data are then compiled into a summary Python dictionary (containing a global frame limit [`nof_values`], a summary map of the pairwise distances [`pairs_data`], and the column index mapping from the original to the filtered [`column_mapping_original_to_filtered`]) which can then be used in subsequent quantitative analysis.

Chapter 4 - Methods: NOE Averaging and Violation Analysis

4.1 Introduction

Once the grouping has been implemented so that each cluster of atoms can be traced to an individual, unambiguous experimental constraint (in other words, the constraints have no degeneracy), the process proceeds to the final part of the work, the quantitative analysis phase. Here the software will read the distance matrix and determine group averages, as described in this chapter, and compare them to the upper limit distance values from the constraints to produce a complete NOE violation list.

4.2 The Theory of NOE Averaging

How intense an NOE signal is, approximately, is proportional to the ensemble average of the inverse sixth power of the inter-proton distance, noted as “ r^{-6} ”. The effective distance of this average can be calculated from the formula: $\{ r^{-6} \}^{-1/6}$

We must then compare this to the upper bound of the restraint, which is experimental data for the distance restraint. The script also computes the $(r^{-3})^{-1/3}$, or the inverse third average. The difference between the two averages gives you a measure of the distribution of the data, as discussed further below. If the difference is large, there is local flexibility/dynamics and the conformations will be spread out (wide conformational ensemble, possibly quite asymmetric).

4.3 The `calculate_noe_sums_in_memory` Function

The steps of the numerical averaging process are embedded within the `calculate_noe_sums_in_memory` function. Note that rather than searching through a text file, the function loops frame by frame through the `carma.distances` data and executes the frame-level degeneracy resolution (`min()`) and the inverse-power accumulations simultaneously.

Code Excerpt 4.1: Running Sum Accumulation in `calculate_noe_sums_in_memory`

```

# Initialize sum arrays and frame counter
sum_r3 = np.zeros(num_unique_groups)
sum_r6 = np.zeros(num_unique_groups)
n_frames = 0

# ... [Iterates through carma_file line by line] ...
n_frames += 1

# For each unique group, find the minimum distance and
add to running sums
for i, group_key in enumerate(unique_restraint_keys):
    raw_indices_for_group = column_groups[group_key]
    try:
        group_distances = [float(distances[j]) for j
in raw_indices_for_group]
    except (ValueError, IndexError) as e:
        continue
    if not group_distances:
        continue

    min_dist = min(group_distances)

# Add to the sums (with an epsilon to prevent
division by zero)
sum_r3[i] += 1.0 / ((min_dist + 1e-9) ** 3)
sum_r6[i] += 1.0 / ((min_dist + 1e-9) ** 6)

```

Explanation: This function is done in large part with a little help from Numpy (Harris, C.R., Millman, K.J., van der Walt, S.J. (2020)), a free open-source Python package for numerical computation. Numpy contains fast and clean C-code under the surface allowing for efficient and optimized multi-dimensional arrays and mathematical functions. The functions above first assign 2 flat Numpy arrays (sum_r3 and sum_r6), each with initial value 0. Then, the number of frames (n_frames) is kept count by a running number during the iteration through each feasible line of the trajectory. Within the loop over each line of the trajectory, the lowest resolved distance (min_dist) in each restraint-group is determined, and that distance is inverted to a 3rd and 6th power, respectively, to calculate the inverse distance of this line at the appropriate place in the array. To avoid numerical errors resulting from a division-by-zero crash (e.g. from atom-centers coinciding at a specific frame), a tiny numerical constant is continuously added to the denominator.

4.4 The `run_noe_averaging` Function: Generating the Final Report

Here the post-processing translation engine starts a series of functions to return the data to a format which is suitable for presentation. This series starts with the function `run_noe_averaging` which collects all of the values of our packed mathematical summation structure. This is a repetitive function that will perform all of the following operations for each experimental datapoint successfully mapped:

- **Ensemble Distance Derivation:** The script computes the final physical time-average metrics from the total accumulated array elements using the target biophysical equations (Daura, X., Antes, I., van Gunsteren, W.F., Thiel, W. and Mark, A.E. (1999)):

- $avg_r3_dist = (\sum r^3 - 3 / Nframes)^{1/3}$
- $avg_r6_dist = (\sum r^6 - 6 / Nframes)^{1/6}$

- **Quantifying Deviations and Violations:**

- Deviation: Calculated explicitly as `avg_r6_dist - exp_distance`. A negative value reveals that the simulated ensemble trajectory remains well within the boundary limits of the experimental constraint, whereas a positive profile signifies an inflation away from the targeted upper limit
- Violation: Defined rigorously as `max(0, deviation)` to filter out compliant assignments and isolate only the exact structural errors exceeding the experimental limit.

- **Structured Terminal Output Rendering:** The function maps structural definitions using an internal dictionary tracker (`column_mapping_original_to_filtered`). It uses a loop through the verified entries and outputs an aligned terminal table. Each row displays the experimental limit, the r^3 and r^6 simulation bounds, the formal deviation metric, and the final violation threshold. If a violation is observed, the value is automatically color-coded using system terminal ANSI escape sequences based on structural severity to organize and condense troubleshooting.

- **Compiling Global Statistics:** To provide an instantaneous overview of global model accuracy, the script calculates summary values at the footer of the output. These include the absolute count and relative percentage of violated restraints, the absolute maximum violation detected, and a stratified frequency distribution profiling errors across strict severity boundaries (e.g. 0–0.5 Å, >0.5 Å, >1.0 Å, etc.). This localized profiling helps users quickly identify mismatched domains, effectively guiding further molecular refinements.

Chapter 5 - Results: Case Study - Analysis of the Chignolin Variant CLN025

As a case study of the proposed framework on a dynamic system, we choose CLN025, a 10-residue variant of chignolin (a former 10-residue peptide developed three years prior by the same group [Honda, S., Yamasaki, K., Sawada, Y., & Morii, H. (2004).]), which is an artificial designed miniprotein with very fast folding rate (Honda et al., 2008). Due to the small size, defined secondary structures, and β -hairpin topology of CLN025 (which consist of the sequence: YYDPETGTWY), it is a useful test for the principle of protein folding, and a standard system to evaluate the force field in molecular simulation. The variant of CLN025 that was engineered (with its attributes granted by the swapping of terminal glycine with tyrosine residues) features greater thermal stability and higher thermal resilience allowing it to be much more easily used as a model system. Since CLN025 consists of only ten amino acids and subsequently around 170 atoms including hydrogen, it allows us to examine its complete folding and unfolding pathway utilizing molecular dynamic simulations in silico in a time-saving manner that would be impossible while proceeding with the same test using larger protein models, at least within the scope of this dissertation. The tyrosine residues mentioned above lead to strong aromatic interactions and thus high structural stability in a β -hairpin turn. With all of the factors above accounted for, we can surmise that if NOEmd's results don't align with the predicted NMR chemical shifts of the protein, certain calibration errors must have been made and the program would be unreliable.

However, comparing the data derived from a previous study with how the program handles the residues below, we can see that the automation workflow results from the script completely track with the predicted results.

Data from previous calculations [Table 5.1] conducted by (Athanasia-Panagiota Serafeim; Georgios Salamanos; Kalliopi K. Patapati; Nicholas M. Glykos, 2016) were used as a baseline to compare and contrast with the NOEmd experimentally derived findings of the second table [Table 5.2].

Going back and forth whilst checking each entry pair manually, it is determined that the output table results align with the previous study's findings, further proving the fail-proof nature of the scripting, especially with how it concerns the parsing function of the program.

Table 5.1: NOEs from the calculations conducted in the 2016 study.

CLN025 : experimental vs MD-derived NOEs and their estimated standard deviations

Proton Indices	NOE (EXP)	NOE (STAR)	σ (STAR)	NOE (ILDN)	σ (ILDN)	NOE (NMR)	σ (NMR)
122-126	3.75	2.611263	0.538095	2.589342	0.494052	2.534155	0.314964
126-146	4.66	3.597264	0.556447	3.685342	0.510994	4.020951	0.316087
108-112	3.61	2.781489	0.484556	2.771055	0.472551	2.742704	0.443714
127-130	3.82	2.93632	0.580237	2.923893	0.577842	2.889541	0.569727
126-130	3.82	3.125396	0.54962	3.143341	0.545991	3.189156	0.534267
108-114	4.51	4.214299	0.544288	4.244267	0.528531	4.358816	0.478145
93-101	4.75	5.064103	0.423772	5.08835	0.415293	5.102678	0.412836
114-146	4.82	5.875258	1.170809	5.824237	1.164016	5.582342	1.154325
30-46	4.82	4.019499	0.273389	4.05379	0.241924	4.034287	0.264627
29-46	4.82	4.02309	0.26051	4.02234	0.269612	4.064283	0.238731
46-50	3.56	3.036333	0.28654	3.009743	0.268291	2.990794	0.236223
46-51	3.56	2.557921	0.257028	2.536876	0.232684	2.512684	0.197046
59-72	3.66	2.980893	0.386405	2.974821	0.379217	2.968205	0.344784
33-65	4.26	6.469092	1.647051	6.483796	1.582267	6.603936	1.556512
33-132	5.02	5.032921	2.489723	4.941119	2.527135	4.716088	2.421841
35-132	4.68	5.288005	2.774603	5.148759	2.871302	4.998808	2.783247
35-139	5.50	5.944403	2.796608	5.862881	2.656843	5.549402	2.272082
6-12	3.81	3.284419	0.819923	3.295452	0.816888	3.504374	0.721132
112-156	4.39	3.63943	2.489036	3.526713	2.257834	3.356214	2.140418
114-156	4.66	4.371134	2.537939	4.282522	2.233512	4.244754	2.160627
27-35	4.51	5.339021	0.599355	5.306223	0.607211	5.232895	0.615931
146-163	3.74	3.591461	1.000734	3.662933	1.004141	3.518086	1.036651
122-130	4.74	3.502382	1.084235	3.550816	1.072898	3.618254	1.065333
101-108	3.44	2.687964	0.701233	2.663091	0.632297	2.663896	0.351343
122-141	5.37	3.365842	1.317315	3.354144	1.321774	3.346444	1.339449
87-101	2.82	2.16851	0.310808	2.156553	0.329216	2.299892	0.29949
72-87	3.29	2.483315	0.301629	2.483952	0.287789	2.66302	0.267832
46-154	4.97	4.838374	2.594991	4.564385	2.664167	4.506604	2.266764

40-46	4.37	4.736105	0.779407	4.770073	0.748001	4.833017	0.739288
122-127	3.75	3.189703	0.463282	3.142397	0.487015	3.1216	0.508849
127-146	4.66	3.146561	0.578307	3.151062	0.568785	3.211784	0.545467
112-122	4.16	3.499187	0.590874	3.521502	0.580679	3.544805	0.566839
112-146	5.02	5.928225	0.589784	5.949238	0.588982	5.770388	0.544264
110-146	5.50	5.964486	0.502432	6.034735	0.427114	6.016305	0.362318
27-46	3.13	2.207062	0.172952	2.198623	0.147315	2.19668	0.151263
110-122	2.94	2.230195	0.292339	2.215435	0.220633	2.213718	0.228635
33-62	4.85	5.825221	1.888518	5.874074	1.828302	5.704595	1.868715
35-65	3.22	5.315469	2.127923	5.339321	1.970031	5.598151	1.959192
27-124	4.10	2.59508	2.084526	2.545076	1.827179	2.47893	1.337663
87-93	4.06	3.542166	0.66625	3.530442	0.656217	3.335784	0.648174
114-122	4.14	3.294734	0.856329	3.280386	0.826915	3.216308	0.808561
6-19	4.74	5.454479	0.684971	5.435217	0.686303	5.695744	0.676185
148-154	3.64	3.285968	0.694244	3.344444	0.675751	3.282436	0.677179
148-156	5.50	5.336914	0.548913	5.377438	0.541925	5.303504	0.509195
14-48	3.96	4.863568	2.748349	4.657779	2.64969	4.696577	2.373711
12-48	4.32	5.787922	1.731603	5.643744	1.690827	5.293448	1.79898
124-141	3.93	3.706897	0.658395	3.688694	0.655497	3.622124	0.655022
124-130	4.55	3.582732	0.573457	3.569232	0.566882	3.52859	0.542416
35-62	5.21	5.63619	2.133514	5.705709	1.997492	5.669815	2.040834
33-68	3.76	4.781491	1.242232	4.784082	1.149071	4.794126	1.13406
35-68	4.78	3.316197	1.803932	3.351003	1.592067	3.348361	1.633349
33-59	5.01	5.46119	1.494116	5.496767	1.451313	5.546737	1.441237
35-124	4.41	6.032474	2.130528	5.910521	2.045829	5.689949	1.664222
27-163	5.50	4.402874	2.441918	4.486989	2.522242	4.453093	2.213683
112-154	4.93	5.180383	1.6504	5.003461	1.543339	4.876362	1.469618
91-108	4.87	4.137942	0.907928	4.101326	0.860539	4.071301	0.596155
124-163	4.94	4.179185	1.479934	4.273717	1.466191	4.270643	1.43551
35-141	4.82	5.706337	2.555284	5.567262	2.446804	5.356828	2.129189
35-130	3.98	5.881998	2.383553	5.744857	2.470663	5.571386	2.317395
33-141	4.45	4.571872	2.588202	4.423927	2.518171	4.190296	2.264338
74-79	4.00	2.806485	0.460675	2.797467	0.462702	2.884604	0.459471
48-59	3.25	2.343479	0.29979	2.331307	0.290712	2.372057	0.27804
19-50	4.73	3.767346	2.93328	3.760237	2.761316	3.663907	2.590303

25-29	3.68	3.077913	0.297201	3.079197	0.29583	2.979677	0.287676
29-59	4.42	7.201415	0.526587	7.173263	0.51028	7.132571	0.458847
29-141	4.53	3.715266	2.537746	3.640735	2.36437	3.434512	2.109242
29-139	4.27	4.157322	2.894416	4.100192	2.723361	3.862507	2.490748
33-126	5.04	4.879706	2.512838	4.698596	2.503227	4.421095	2.328356
46-50	3.10	3.036333	0.28654	3.009743	0.268291	2.990794	0.236223
50-72	4.62	4.582001	0.473684	4.657972	0.405883	4.741895	0.401063
50-91	4.52	5.629558	1.056132	5.792717	0.898546	6.164707	0.63425
50-108	3.61	4.729146	1.350505	4.64359	1.17846	4.523274	0.73378
50-112	4.33	4.501912	2.134667	4.398521	1.907666	4.327387	1.377877
50-154	4.66	4.607829	2.916911	4.331876	2.846053	4.175363	2.258976
50-156	3.76	3.717754	3.418357	3.501133	3.290203	3.423573	2.591413
122-126	3.04	2.611263	0.538095	2.589342	0.494052	2.534155	0.314964
126-130	3.33	3.125396	0.54962	3.143341	0.545991	3.189156	0.534267
126-141	3.44	3.028881	0.724638	3.011334	0.720955	2.972841	0.706843
126-146	4.03	3.597264	0.556447	3.685342	0.510994	4.020951	0.316087
146-150	3.54	2.675047	0.424641	2.676857	0.434719	2.575172	0.339306

Table 5.2: Results derived from the NOEmd analysis, presented each time the first output is printed when the run concludes.

```

=====
122 126 *** 122 A 9 TRP H <==> 126 A 9 TRP HB2
126 146 *** 126 A 9 TRP HB2 <==> 146 A 10 TYR H
108 112 *** 108 A 8 THR H <==> 112 A 8 THR HB
127 130 *** 127 A 9 TRP HB3 <==> 130 A 9 TRP HD1
126 130 *** 126 A 9 TRP HB2 <==> 130 A 9 TRP HD1
108 114 *** 108 A 8 THR H <==> 114 A 8 THR QG2 (Degenerate)
108 115 *** 108 A 8 THR H <==> 115 A 8 THR QG2 (Degenerate)
108 116 *** 108 A 8 THR H <==> 116 A 8 THR QG2 (Degenerate)
93 101 *** 93 A 6 THR QG2 <==> 101 A 7 GLY H (Degenerate)
94 101 *** 94 A 6 THR QG2 <==> 101 A 7 GLY H (Degenerate)
95 101 *** 95 A 6 THR QG2 <==> 101 A 7 GLY H (Degenerate)
114 146 *** 114 A 8 THR QG2 <==> 146 A 10 TYR H (Degenerate)

```

115	146	***	115 A	8 THR	QG2 <==>	146 A	10 TYR	H (Degenerate)
116	146	***	116 A	8 THR	QG2 <==>	146 A	10 TYR	H (Degenerate)
30	46	***	30 A	2 TYR	HB3 <==>	46 A	3 ASP	H
29	46	***	29 A	2 TYR	HB2 <==>	46 A	3 ASP	H
46	50	***	46 A	3 ASP	H <==>	50 A	3 ASP	HB2
46	51	***	46 A	3 ASP	H <==>	51 A	3 ASP	HB3
59	72	***	59 A	4 PRO	QD <==>	72 A	5 GLU	H (Degenerate)
60	72	***	60 A	4 PRO	QD <==>	72 A	5 GLU	H (Degenerate)
33	65	***	33 A	2 TYR	QD <==>	65 A	4 PRO	QB (Degenerate)
33	66	***	33 A	2 TYR	QD <==>	66 A	4 PRO	QB (Degenerate)
42	65	***	42 A	2 TYR	QD <==>	65 A	4 PRO	QB (Degenerate)
42	66	***	42 A	2 TYR	QD <==>	66 A	4 PRO	QB (Degenerate)
33	132	***	33 A	2 TYR	QD <==>	132 A	9 TRP	HE1 (Degenerate)
42	132	***	42 A	2 TYR	QD <==>	132 A	9 TRP	HE1 (Degenerate)
35	132	***	35 A	2 TYR	QE <==>	132 A	9 TRP	HE1 (Degenerate)
40	132	***	40 A	2 TYR	QE <==>	132 A	9 TRP	HE1 (Degenerate)
35	139	***	35 A	2 TYR	QE <==>	139 A	9 TRP	HZ3 (Degenerate)
40	139	***	40 A	2 TYR	QE <==>	139 A	9 TRP	HZ3 (Degenerate)
6	12	***	6 A	1 TYR	HA <==>	12 A	1 TYR	QD (Degenerate)
6	21	***	6 A	1 TYR	HA <==>	21 A	1 TYR	QD (Degenerate)
112	156	***	112 A	8 THR	HB <==>	156 A	10 TYR	QE (Degenerate)
112	161	***	112 A	8 THR	HB <==>	161 A	10 TYR	QE (Degenerate)
114	156	***	114 A	8 THR	QG2 <==>	156 A	10 TYR	QE (Degenerate)
114	161	***	114 A	8 THR	QG2 <==>	161 A	10 TYR	QE (Degenerate)
115	156	***	115 A	8 THR	QG2 <==>	156 A	10 TYR	QE (Degenerate)
115	161	***	115 A	8 THR	QG2 <==>	161 A	10 TYR	QE (Degenerate)
116	156	***	116 A	8 THR	QG2 <==>	156 A	10 TYR	QE (Degenerate)
116	161	***	116 A	8 THR	QG2 <==>	161 A	10 TYR	QE (Degenerate)
27	35	***	27 A	2 TYR	HA <==>	35 A	2 TYR	QE (Degenerate)

27	40	***	27 A	2 TYR	HA <==>	40 A	2 TYR	QE (Degenerate)
146	154	***	146 A	10 TYR	H <==>	154 A	10 TYR	QD (Degenerate)
146	163	***	146 A	10 TYR	H <==>	163 A	10 TYR	QD (Degenerate)
122	130	***	122 A	9 TRP	H <==>	130 A	9 TRP	HD1
101	108	***	101 A	7 GLY	H <==>	108 A	8 THR	H
122	141	***	122 A	9 TRP	H <==>	141 A	9 TRP	HE3
87	101	***	87 A	6 THR	H <==>	101 A	7 GLY	H
72	87	***	72 A	5 GLU	H <==>	87 A	6 THR	H
46	154	***	46 A	3 ASP	H <==>	154 A	10 TYR	QD (Degenerate)
46	163	***	46 A	3 ASP	H <==>	163 A	10 TYR	QD (Degenerate)
35	46	***	35 A	2 TYR	QE <==>	46 A	3 ASP	H (Degenerate)
40	46	***	40 A	2 TYR	QE <==>	46 A	3 ASP	H (Degenerate)
122	127	***	122 A	9 TRP	H <==>	127 A	9 TRP	HB3
127	146	***	127 A	9 TRP	HB3 <==>	146 A	10 TYR	H
112	122	***	112 A	8 THR	HB <==>	122 A	9 TRP	H
112	146	***	112 A	8 THR	HB <==>	146 A	10 TYR	H
110	146	***	110 A	8 THR	HA <==>	146 A	10 TYR	H
27	46	***	27 A	2 TYR	HA <==>	46 A	3 ASP	H
110	122	***	110 A	8 THR	HA <==>	122 A	9 TRP	H
33	62	***	33 A	2 TYR	QD <==>	62 A	4 PRO	QG (Degenerate)
33	63	***	33 A	2 TYR	QD <==>	63 A	4 PRO	QG (Degenerate)
42	62	***	42 A	2 TYR	QD <==>	62 A	4 PRO	QG (Degenerate)
42	63	***	42 A	2 TYR	QD <==>	63 A	4 PRO	QG (Degenerate)
35	65	***	35 A	2 TYR	QE <==>	65 A	4 PRO	QB (Degenerate)
35	66	***	35 A	2 TYR	QE <==>	66 A	4 PRO	QB (Degenerate)
40	65	***	40 A	2 TYR	QE <==>	65 A	4 PRO	QB (Degenerate)
40	66	***	40 A	2 TYR	QE <==>	66 A	4 PRO	QB (Degenerate)
27	124	***	27 A	2 TYR	HA <==>	124 A	9 TRP	HA
87	93	***	87 A	6 THR	H <==>	93 A	6 THR	QG2 (Degenerate)

87	94	***	87 A	6 THR	H <==>	94 A	6 THR	QG2 (Degenerate)
87	95	***	87 A	6 THR	H <==>	95 A	6 THR	QG2 (Degenerate)
114	122	***	114 A	8 THR	QG2 <==>	122 A	9 TRP	H (Degenerate)
115	122	***	115 A	8 THR	QG2 <==>	122 A	9 TRP	H (Degenerate)
116	122	***	116 A	8 THR	QG2 <==>	122 A	9 TRP	H (Degenerate)
6	14	***	6 A	1 TYR	HA <==>	14 A	1 TYR	QE (Degenerate)
6	19	***	6 A	1 TYR	HA <==>	19 A	1 TYR	QE (Degenerate)
148	154	***	148 A	10 TYR	HA <==>	154 A	10 TYR	QD (Degenerate)
148	163	***	148 A	10 TYR	HA <==>	163 A	10 TYR	QD (Degenerate)
148	156	***	148 A	10 TYR	HA <==>	156 A	10 TYR	QE (Degenerate)
148	161	***	148 A	10 TYR	HA <==>	161 A	10 TYR	QE (Degenerate)
14	48	***	14 A	1 TYR	QE <==>	48 A	3 ASP	HA (Degenerate)
19	48	***	19 A	1 TYR	QE <==>	48 A	3 ASP	HA (Degenerate)
12	48	***	12 A	1 TYR	QD <==>	48 A	3 ASP	HA (Degenerate)
21	48	***	21 A	1 TYR	QD <==>	48 A	3 ASP	HA (Degenerate)
124	141	***	124 A	9 TRP	HA <==>	141 A	9 TRP	HE3
124	130	***	124 A	9 TRP	HA <==>	130 A	9 TRP	HD1
35	62	***	35 A	2 TYR	QE <==>	62 A	4 PRO	QG (Degenerate)
35	63	***	35 A	2 TYR	QE <==>	63 A	4 PRO	QG (Degenerate)
40	62	***	40 A	2 TYR	QE <==>	62 A	4 PRO	QG (Degenerate)
40	63	***	40 A	2 TYR	QE <==>	63 A	4 PRO	QG (Degenerate)
33	68	***	33 A	2 TYR	QD <==>	68 A	4 PRO	HA (Degenerate)
42	68	***	42 A	2 TYR	QD <==>	68 A	4 PRO	HA (Degenerate)
35	68	***	35 A	2 TYR	QE <==>	68 A	4 PRO	HA (Degenerate)
40	68	***	40 A	2 TYR	QE <==>	68 A	4 PRO	HA (Degenerate)
33	59	***	33 A	2 TYR	QD <==>	59 A	4 PRO	QD (Degenerate)
33	60	***	33 A	2 TYR	QD <==>	60 A	4 PRO	QD (Degenerate)
42	59	***	42 A	2 TYR	QD <==>	59 A	4 PRO	QD (Degenerate)
42	60	***	42 A	2 TYR	QD <==>	60 A	4 PRO	QD (Degenerate)

35	124	***	35 A	2 TYR	QE <==>	124 A	9 TRP	HA (Degenerate)
40	124	***	40 A	2 TYR	QE <==>	124 A	9 TRP	HA (Degenerate)
27	154	***	27 A	2 TYR	HA <==>	154 A	10 TYR	QD (Degenerate)
27	163	***	27 A	2 TYR	HA <==>	163 A	10 TYR	QD (Degenerate)
112	154	***	112 A	8 THR	HB <==>	154 A	10 TYR	QD (Degenerate)
112	163	***	112 A	8 THR	HB <==>	163 A	10 TYR	QD (Degenerate)
91	108	***	91 A	6 THR	HB <==>	108 A	8 THR	H
124	154	***	124 A	9 TRP	HA <==>	154 A	10 TYR	QD (Degenerate)
124	163	***	124 A	9 TRP	HA <==>	163 A	10 TYR	QD (Degenerate)
35	141	***	35 A	2 TYR	QE <==>	141 A	9 TRP	HE3 (Degenerate)
40	141	***	40 A	2 TYR	QE <==>	141 A	9 TRP	HE3 (Degenerate)
35	130	***	35 A	2 TYR	QE <==>	130 A	9 TRP	HD1 (Degenerate)
40	130	***	40 A	2 TYR	QE <==>	130 A	9 TRP	HD1 (Degenerate)
33	141	***	33 A	2 TYR	QD <==>	141 A	9 TRP	HE3 (Degenerate)
42	141	***	42 A	2 TYR	QD <==>	141 A	9 TRP	HE3 (Degenerate)
74	79	***	74 A	5 GLU	HA <==>	79 A	5 GLU	QG (Degenerate)
74	80	***	74 A	5 GLU	HA <==>	80 A	5 GLU	QG (Degenerate)
48	59	***	48 A	3 ASP	HA <==>	59 A	4 PRO	QD (Degenerate)
48	60	***	48 A	3 ASP	HA <==>	60 A	4 PRO	QD (Degenerate)
14	50	***	14 A	1 TYR	QE <==>	50 A	3 ASP	QB (Degenerate)
14	51	***	14 A	1 TYR	QE <==>	51 A	3 ASP	QB (Degenerate)
19	50	***	19 A	1 TYR	QE <==>	50 A	3 ASP	QB (Degenerate)
19	51	***	19 A	1 TYR	QE <==>	51 A	3 ASP	QB (Degenerate)
25	29	***	25 A	2 TYR	H <==>	29 A	2 TYR	QB (Degenerate)
25	30	***	25 A	2 TYR	H <==>	30 A	2 TYR	QB (Degenerate)
29	59	***	29 A	2 TYR	QB <==>	59 A	4 PRO	QD (Degenerate)
29	60	***	29 A	2 TYR	QB <==>	60 A	4 PRO	QD (Degenerate)
30	59	***	30 A	2 TYR	QB <==>	59 A	4 PRO	QD (Degenerate)
30	60	***	30 A	2 TYR	QB <==>	60 A	4 PRO	QD (Degenerate)

29	141	***	29 A	2 TYR	QB <==>	141 A	9 TRP	HE3 (Degenerate)
30	141	***	30 A	2 TYR	QB <==>	141 A	9 TRP	HE3 (Degenerate)
29	139	***	29 A	2 TYR	QB <==>	139 A	9 TRP	HZ3 (Degenerate)
30	139	***	30 A	2 TYR	QB <==>	139 A	9 TRP	HZ3 (Degenerate)
33	126	***	33 A	2 TYR	QD <==>	126 A	9 TRP	QB (Degenerate)
33	127	***	33 A	2 TYR	QD <==>	127 A	9 TRP	QB (Degenerate)
42	126	***	42 A	2 TYR	QD <==>	126 A	9 TRP	QB (Degenerate)
42	127	***	42 A	2 TYR	QD <==>	127 A	9 TRP	QB (Degenerate)
50	72	***	50 A	3 ASP	QB <==>	72 A	5 GLU	H (Degenerate)
51	72	***	51 A	3 ASP	QB <==>	72 A	5 GLU	H (Degenerate)
50	91	***	50 A	3 ASP	QB <==>	91 A	6 THR	HB (Degenerate)
51	91	***	51 A	3 ASP	QB <==>	91 A	6 THR	HB (Degenerate)
50	108	***	50 A	3 ASP	QB <==>	108 A	8 THR	H (Degenerate)
51	108	***	51 A	3 ASP	QB <==>	108 A	8 THR	H (Degenerate)
50	112	***	50 A	3 ASP	QB <==>	112 A	8 THR	HB (Degenerate)
51	112	***	51 A	3 ASP	QB <==>	112 A	8 THR	HB (Degenerate)
50	154	***	50 A	3 ASP	QB <==>	154 A	10 TYR	QD (Degenerate)
50	163	***	50 A	3 ASP	QB <==>	163 A	10 TYR	QD (Degenerate)
51	154	***	51 A	3 ASP	QB <==>	154 A	10 TYR	QD (Degenerate)
51	163	***	51 A	3 ASP	QB <==>	163 A	10 TYR	QD (Degenerate)
50	156	***	50 A	3 ASP	QB <==>	156 A	10 TYR	QE (Degenerate)
50	161	***	50 A	3 ASP	QB <==>	161 A	10 TYR	QE (Degenerate)
51	156	***	51 A	3 ASP	QB <==>	156 A	10 TYR	QE (Degenerate)
51	161	***	51 A	3 ASP	QB <==>	161 A	10 TYR	QE (Degenerate)
126	141	***	126 A	9 TRP	QB <==>	141 A	9 TRP	HE3 (Degenerate)
127	141	***	127 A	9 TRP	QB <==>	141 A	9 TRP	HE3 (Degenerate)
146	150	***	146 A	10 TYR	H <==>	150 A	10 TYR	QB (Degenerate)
146	151	***	146 A	10 TYR	H <==>	151 A	10 TYR	QB (Degenerate)

=====

Comparative Result Analysis: CLN025 Trajectories

This analysis evaluates the structural performance of two molecular dynamics simulation trajectories: the ILDN Trajectory ([cln025_adapt_ILDN_COMPLETE.dcd](#)) and the STAR-ILDN Trajectory ([cln025_adapt_STAR_ILDN_COMPLETE.dcd](#)). Performance is measured against Nuclear Overhauser Effect (NOE) experimental distance restraints, which serve as metrics for structural accuracy in protein folding simulations.

Summary Statistics

A direct comparison of the global NOE violation profiles reveals that both frameworks accurately capture the overall conformational complex of CLN025, though they exhibit subtle differences in their error distributions:

Metric	ILDN Trajectory	STAR Trajectory	Difference / Shift
Total NOE Restraints	76	76	0 (Identical dataset)
Violated Restraints (>0.001 Å)	7 (9.2%)	8 (10.5%)	+1 violation in STAR
Average Violation Magnitude	0.318 Å	0.307 Å	-0.011 Å (Slightly lower in STAR)
Maximum Violation	0.436 Å	0.516 Å	+0.080 Å (Higher peak error in STAR)
Severity: 0–0.5 Å	7	7	0
Severity: >0.5 Å	0	1	+1 in STAR

Table 5.3 : Presentation of the shifts between the ILDN and STAR-ILDN trajectories.

Key Structural Differences and Variations

While the majority of the sub-angstrom distance averages (r^3) and (r^6) match very closely between the two datasets, specific regions highlight structural divergence:

1. The Core Divergent Anchor: Tyr2 – Pro4

Both datasets explicitly flag the interaction between **2 TYR QD** and **4 PRO HA** as a primary "Divergent" zone.

- **ILDN Trajectory:** Deviation of 0.430 Å (with an ensemble distance r^6 4.190 Å versus an EXP of 3.76 Å).
- **STAR-ILDN Trajectory:** The violation increases to **0.516 Å** (r^6 = 4.276 Å), pushing this specific restraint into the 0.5 Å severity bracket.

In the meantime, while examining the 2 TYR QD $\rightleftharpoons$ 4 PRO QB neighbor interaction:

- The ILDN Trajectory manages to satisfy this without violation (Deviation: -0.037 Å).
- The STAR-ILDN Trajectory appears to have a small violation of **0.048 Å**.

2. Long-Range Aromatic Packing: Tyr2 – Trp9

The tertiary hydrophobic core of CLN025 is stabilized by aromatic stacking interactions involving Tyr2 and Trp9, and while each trajectory incurs relatively small violations in this respect, the distributions differ:

- **2 TYR QE <==> 9 TRP HD1**: Produced a violation of 0.435 Å in ILDN, and slightly higher at **0.472 Å** in STAR-ILDN.
- **2 TYR QE <==> 9 TRP HA**: Produced a violation of 0.302 Å in ILDN, but appeared lower at **0.261 Å** in STAR-ILDN.

3. Backbone Terminal Regions

Interactions involving the termini (such as **1 TYR QD <==> 3 ASP HA**) show a slight increase in violation in the STAR-ILDN trajectory (0.479 Å) compared to the ILDN trajectory (0.436 Å).

The complete trajectory tables are formatted below with each divergent pair emphasized in **bold and underlined** text next to each line for clear identification.

➤ ILDN Trajectory of CLN025 (cln025_adapt_ILDN_COMPLETE.dcd):

NOE Analysis Results:

EXP	<r ⁻³ >	<r ⁻⁶ >	Deviation	Violation	Atoms Involved			
3.75	2.692	2.601	-1.149	0.000	9 TRP	H <==>	9 TRP	HB2
4.66	3.566	3.275	-1.385	0.000	9 TRP	HB2 <==>	10 TYR	H
3.61	2.935	2.825	-0.785	0.000	8 THR	H <==>	8 THR	HB
3.82	2.988	2.872	-0.948	0.000	9 TRP	HB3 <==>	9 TRP	HD1
3.82	3.104	2.969	-0.851	0.000	9 TRP	HB2 <==>	9 TRP	HD1
4.51	3.338	3.049	-1.461	0.000	8 THR	H <==>	8 THR	QG2
4.75	4.530	4.271	-0.479	0.000	6 THR	QG2 <==>	7 GLY	H
4.82	5.007	4.194	-0.626	0.000	8 THR	QG2 <==>	10 TYR	H
4.82	3.959	3.759	-1.061	0.000	2 TYR	HB3 <==>	3 ASP	H
4.82	3.934	3.711	-1.109	0.000	2 TYR	HB2 <==>	3 ASP	H

3.56	3.061	2.986	-0.574	0.000	3 ASP	H <==>	3 ASP	HB2	Divergent
3.56	2.666	2.609	-0.951	0.000	3 ASP	H <==>	3 ASP	HB3	
3.66	3.014	2.922	-0.738	0.000	4 PRO	QD <==>	5 GLU	H	
4.26	4.688	4.223	-0.037	0.000	2 TYR	QD <==>	4 PRO	QB	
5.02	4.741	4.129	-0.891	0.000	2 TYR	QD <==>	9 TRP	HE1	
4.68	4.965	4.093	-0.587	0.000	2 TYR	QE <==>	9 TRP	HE1	
5.50	5.327	4.410	-1.090	0.000	2 TYR	QE <==>	9 TRP	HZ3	
3.81	2.758	2.631	-1.179	0.000	1 TYR	HA <==>	1 TYR	QD	
4.39	3.533	3.043	-1.347	0.000	8 THR	HB <==>	10 TYR	QE	
4.66	3.830	3.372	-1.288	0.000	8 THR	QG2 <==>	10 TYR	QE	
4.51	4.878	4.841	0.331	0.331	2 TYR	HA <==>	2 TYR	QE	
3.74	3.170	2.999	-0.741	0.000	10 TYR	H <==>	10 TYR	QD	
4.74	3.537	3.150	-1.590	0.000	9 TRP	H <==>	9 TRP	HD1	
3.44	2.769	2.653	-0.787	0.000	7 GLY	H <==>	8 THR	H	
5.37	3.525	3.027	-2.343	0.000	9 TRP	H <==>	9 TRP	HE3	
2.82	2.238	2.164	-0.656	0.000	6 THR	H <==>	7 GLY	H	
3.29	2.455	2.383	-0.907	0.000	5 GLU	H <==>	6 THR	H	
4.97	4.447	3.958	-1.012	0.000	3 ASP	H <==>	10 TYR	QD	
4.37	4.718	4.502	0.132	0.132	2 TYR	QE <==>	3 ASP	H	
3.75	3.085	2.939	-0.811	0.000	9 TRP	H <==>	9 TRP	HB3	
4.66	3.284	3.088	-1.572	0.000	9 TRP	HB3 <==>	10 TYR	H	
4.16	3.424	3.111	-1.049	0.000	8 THR	HB <==>	9 TRP	H	
5.02	5.326	4.448	-0.572	0.000	8 THR	HB <==>	10 TYR	H	
5.50	5.281	4.817	-0.683	0.000	8 THR	HA <==>	10 TYR	H	
3.13	2.282	2.245	-0.885	0.000	2 TYR	HA <==>	3 ASP	H	
2.94	2.338	2.281	-0.659	0.000	8 THR	HA <==>	9 TRP	H	
4.85	3.962	3.445	-1.405	0.000	2 TYR	QD <==>	4 PRO	QG	
3.22	3.395	3.012	-0.208	0.000	2 TYR	QE <==>	4 PRO	QB	
4.10	2.843	2.622	-1.478	0.000	2 TYR	HA <==>	9 TRP	HA	
4.06	3.023	2.893	-1.167	0.000	6 THR	H <==>	6 THR	QG2	Divergent
4.14	3.103	2.875	-1.265	0.000	8 THR	QG2 <==>	9 TRP	H	
4.74	4.975	4.898	0.158	0.158	1 TYR	HA <==>	1 TYR	QE	
3.64	2.811	2.697	-0.943	0.000	10 TYR	HA <==>	10 TYR	QD	
5.50	4.990	4.936	-0.564	0.000	10 TYR	HA <==>	10 TYR	QE	
3.96	4.422	3.779	-0.181	0.000	1 TYR	QE <==>	3 ASP	HA	
4.32	5.220	4.756	0.436	0.436	1 TYR	QD <==>	3 ASP	HA	
3.93	3.615	3.296	-0.634	0.000	9 TRP	HA <==>	9 TRP	HE3	
4.55	3.512	3.267	-1.283	0.000	9 TRP	HA <==>	9 TRP	HD1	
5.21	3.617	3.129	-2.081	0.000	2 TYR	QE <==>	4 PRO	QG	
3.76	4.530	4.190	0.430	0.430	2 TYR	QD <==>	4 PRO	HA	
4.78	3.190	2.900	-1.880	0.000	2 TYR	QE <==>	4 PRO	HA	
5.01	3.417	3.121	-1.889	0.000	2 TYR	QD <==>	4 PRO	QD	
4.41	5.298	4.712	0.302	0.302	2 TYR	QE <==>	9 TRP	HA	
5.50	4.257	3.748	-1.752	0.000	2 TYR	HA <==>	10 TYR	QD	
4.93	4.781	4.314	-0.616	0.000	8 THR	HB <==>	10 TYR	QD	
4.87	4.104	3.796	-1.074	0.000	6 THR	HB <==>	8 THR	H	
4.94	3.879	3.615	-1.325	0.000	9 TRP	HA <==>	10 TYR	QD	
4.82	4.954	4.165	-0.655	0.000	2 TYR	QE <==>	9 TRP	HE3	
3.98	5.286	4.415	0.435	0.435	2 TYR	QE <==>	9 TRP	HD1	
4.45	4.008	3.418	-1.032	0.000	2 TYR	QD <==>	9 TRP	HE3	
4.00	2.611	2.553	-1.447	0.000	5 GLU	HA <==>	5 GLU	QG	
3.25	2.184	2.155	-1.095	0.000	3 ASP	HA <==>	4 PRO	QD	
4.73	3.454	3.008	-1.722	0.000	1 TYR	QE <==>	3 ASP	QB	
3.68	2.663	2.604	-1.076	0.000	2 TYR	H <==>	2 TYR	QB	
4.42	4.689	4.379	-0.041	0.000	2 TYR	QB <==>	4 PRO	QD	
4.53	4.086	3.353	-1.177	0.000	2 TYR	QB <==>	9 TRP	HE3	
4.27	4.615	3.765	-0.505	0.000	2 TYR	QB <==>	9 TRP	HZ3	
5.04	3.863	3.289	-1.751	0.000	2 TYR	QD <==>	9 TRP	QB	
4.62	3.742	3.385	-1.235	0.000	3 ASP	QB <==>	5 GLU	H	
4.52	4.081	3.350	-1.170	0.000	3 ASP	QB <==>	6 THR	HB	
3.61	3.237	2.906	-0.704	0.000	3 ASP	QB <==>	8 THR	H	
4.33	3.607	3.106	-1.224	0.000	3 ASP	QB <==>	8 THR	HB	
4.66	4.100	3.548	-1.112	0.000	3 ASP	QB <==>	10 TYR	QD	

3.76	3.112	2.722	-1.038	0.000	3 ASP	QB <==>	10 TYR	QE
3.44	2.665	2.635	-0.805	0.000	9 TRP	QB <==>	9 TRP	HE3
3.54	2.658	2.606	-0.934	0.000	10 TYR	H <==>	10 TYR	QB

=====

SUMMARY: 76 unique NOE restraints displayed.

Violations (>0.001Å): 7 (9.2%)

Average violation (for violated restraints): 0.318 Å

Maximum violation: 0.436 Å

Violation severity distribution:

0-0.5Å: 7 restraints

➤ STAR-ILDN Trajectory of CLN025 (cln025_adapt_STAR_ILDN_COMPLETE.dcd):

NOE Analysis Results:

EXP	<r ³ >	<r ⁶ >	Deviation	Violation	Atoms Involved			
3.75	2.719	2.620	-1.130	0.000	9 TRP	H <==>	9 TRP	HB2
4.66	3.519	3.227	-1.433	0.000	9 TRP	HB2 <==>	10 TYR	H
3.61	2.977	2.859	-0.751	0.000	8 THR	H <==>	8 THR	HB
3.82	3.000	2.882	-0.938	0.000	9 TRP	HB3 <==>	9 TRP	HD1
3.82	3.094	2.961	-0.859	0.000	9 TRP	HB2 <==>	9 TRP	HD1
4.51	3.270	2.995	-1.515	0.000	8 THR	H <==>	8 THR	QG2
4.75	4.514	4.277	-0.473	0.000	6 THR	QG2 <==>	7 GLY	H
4.82	5.095	4.253	-0.567	0.000	8 THR	QG2 <==>	10 TYR	H
4.82	3.880	3.620	-1.200	0.000	2 TYR	HB3 <==>	3 ASP	H
4.82	3.899	3.659	-1.161	0.000	2 TYR	HB2 <==>	3 ASP	H
3.56	3.082	2.999	-0.561	0.000	3 ASP	H <==>	3 ASP	HB2
3.56	2.702	2.638	-0.922	0.000	3 ASP	H <==>	3 ASP	HB3
3.66	3.019	2.925	-0.735	0.000	4 PRO	QD <==>	5 GLU	H
4.26	4.832	4.308	0.048	0.048	2 TYR	QD <==>	4 PRO	QB
5.02	4.862	4.191	-0.829	0.000	2 TYR	QD <==>	9 TRP	HE1
4.68	5.053	4.125	-0.555	0.000	2 TYR	QE <==>	9 TRP	HE1
5.50	5.472	4.465	-1.035	0.000	2 TYR	QE <==>	9 TRP	HZ3
3.81	2.755	2.627	-1.183	0.000	1 TYR	HA <==>	1 TYR	QD
4.39	3.638	3.089	-1.301	0.000	8 THR	HB <==>	10 TYR	QE
4.66	3.952	3.432	-1.228	0.000	8 THR	QG2 <==>	10 TYR	QE
4.51	4.883	4.843	0.333	0.333	2 TYR	HA <==>	2 TYR	QE
3.74	3.191	3.014	-0.726	0.000	10 TYR	H <==>	10 TYR	QD
4.74	3.489	3.107	-1.633	0.000	9 TRP	H <==>	9 TRP	HD1
3.44	2.800	2.673	-0.767	0.000	7 GLY	H <==>	8 THR	H
5.37	3.530	3.039	-2.331	0.000	9 TRP	H <==>	9 TRP	HE3
2.82	2.254	2.179	-0.641	0.000	6 THR	H <==>	7 GLY	H
3.29	2.458	2.385	-0.905	0.000	5 GLU	H <==>	6 THR	H
4.97	4.585	4.024	-0.946	0.000	3 ASP	H <==>	10 TYR	QD
4.37	4.776	4.560	0.190	0.190	2 TYR	QE <==>	3 ASP	H
3.75	3.094	2.948	-0.802	0.000	9 TRP	H <==>	9 TRP	HB3
4.66	3.293	3.087	-1.573	0.000	9 TRP	HB3 <==>	10 TYR	H
4.16	3.415	3.104	-1.056	0.000	8 THR	HB <==>	9 TRP	H
5.02	5.271	4.391	-0.629	0.000	8 THR	HB <==>	10 TYR	H
5.50	5.104	4.658	-0.842	0.000	8 THR	HA <==>	10 TYR	H
3.13	2.315	2.270	-0.860	0.000	2 TYR	HA <==>	3 ASP	H
2.94	2.391	2.316	-0.624	0.000	8 THR	HA <==>	9 TRP	H
4.85	4.055	3.491	-1.359	0.000	2 TYR	QD <==>	4 PRO	QG

Divergent

3.22	3.510	3.069	-0.151	0.000	2 TYR	QE <==>	4 PRO	QB	
4.10	2.953	2.676	-1.424	0.000	2 TYR	HA <==>	9 TRP	HA	
4.06	3.016	2.883	-1.177	0.000	6 THR	H <==>	6 THR	QG2	
4.14	3.180	2.926	-1.214	0.000	8 THR	QG2 <==>	9 TRP	H	
4.74	4.973	4.896	0.156	0.156	1 TYR	HA <==>	1 TYR	QE	
3.64	2.796	2.683	-0.957	0.000	10 TYR	HA <==>	10 TYR	QD	
5.50	4.980	4.926	-0.574	0.000	10 TYR	HA <==>	10 TYR	QE	
3.96	4.563	3.874	-0.086	0.000	1 TYR	QE <==>	3 ASP	HA	
4.32	5.295	4.799	0.479	0.479	1 TYR	QD <==>	3 ASP	HA	
3.93	3.628	3.292	-0.638	0.000	9 TRP	HA <==>	9 TRP	HE3	
4.55	3.515	3.261	-1.289	0.000	9 TRP	HA <==>	9 TRP	HD1	
5.21	3.690	3.159	-2.051	0.000	2 TYR	QE <==>	4 PRO	QG	
3.76	4.662	4.276	0.516	0.516	2 TYR	QD <==>	4 PRO	HA	Divergent
4.78	3.284	2.945	-1.835	0.000	2 TYR	QE <==>	4 PRO	HA	
5.01	3.491	3.160	-1.850	0.000	2 TYR	QD <==>	4 PRO	QD	
4.41	5.358	4.671	0.261	0.261	2 TYR	QE <==>	9 TRP	HA	
5.50	4.381	3.807	-1.693	0.000	2 TYR	HA <==>	10 TYR	QD	
4.93	4.917	4.403	-0.527	0.000	8 THR	HB <==>	10 TYR	QD	
4.87	4.154	3.819	-1.051	0.000	6 THR	HB <==>	8 THR	H	
4.94	3.947	3.656	-1.284	0.000	9 TRP	HA <==>	10 TYR	QD	
4.82	5.111	4.238	-0.582	0.000	2 TYR	QE <==>	9 TRP	HE3	
3.98	5.384	4.452	0.472	0.472	2 TYR	QE <==>	9 TRP	HD1	
4.45	4.171	3.507	-0.943	0.000	2 TYR	QD <==>	9 TRP	HE3	
4.00	2.618	2.558	-1.442	0.000	5 GLU	HA <==>	5 GLU	QG	
3.25	2.199	2.173	-1.077	0.000	3 ASP	HA <==>	4 PRO	QD	
4.73	3.566	3.063	-1.667	0.000	1 TYR	QE <==>	3 ASP	QB	
3.68	2.677	2.616	-1.064	0.000	2 TYR	H <==>	2 TYR	QB	
4.42	4.667	4.199	-0.221	0.000	2 TYR	QB <==>	4 PRO	QD	
4.53	4.251	3.436	-1.094	0.000	2 TYR	QB <==>	9 TRP	HE3	
4.27	4.777	3.842	-0.428	0.000	2 TYR	QB <==>	9 TRP	HZ3	
5.04	4.048	3.393	-1.647	0.000	2 TYR	QD <==>	9 TRP	QB	
4.62	3.677	3.318	-1.302	0.000	3 ASP	QB <==>	5 GLU	H	
4.52	4.102	3.337	-1.183	0.000	3 ASP	QB <==>	6 THR	HB	
3.61	3.375	2.984	-0.626	0.000	3 ASP	QB <==>	8 THR	H	
4.33	3.767	3.185	-1.145	0.000	3 ASP	QB <==>	8 THR	HB	
4.66	4.248	3.621	-1.039	0.000	3 ASP	QB <==>	10 TYR	QD	
3.76	3.224	2.773	-0.987	0.000	3 ASP	QB <==>	10 TYR	QE	
3.44	2.666	2.637	-0.803	0.000	9 TRP	QB <==>	9 TRP	HE3	
3.54	2.652	2.600	-0.940	0.000	10 TYR	H <==>	10 TYR	QB	

```

=====
SUMMARY: 76 unique NOE restraints displayed.
Violations (>0.001Å): 8 (10.5%)
Average violation (for violated restraints): 0.307 Å
Maximum violation: 0.516 Å
Violation severity distribution:
  0-0.5Å: 7 restraints
  >0.5Å: 1 restraints

```

Chapter 6 - Results: Case Study 2 - Analysis of 2N0X

2N0X (Zirafi, O., Kim, K. A., Ständker, L., Mohr, K. B., Sauter, D., Heigele, A. & Münch, J. (2015)) corresponds to the 3D NMR solution structure of EPI-X4 (Endogenous Peptide Inhibitor of CXCR4), a 16-residue (~1.83 kDa) fragment derived from human serum albumin. It is derived from an experimental NMR solution method and its primary function is to act as an antagonist for the CXCR4 chemokine receptor (a class-A GPCR) which regulates cell migration, homing and embryonic development. EPI-X4 has the key structural feature of an unusual lasso-like conformational loop which is essential for selective receptor binding, where the basic N-terminal acts as the primary recognition motif, while the C-terminal plays the role of a stabilizer for the fold.

File Name	# Frames Processed	Violated Restraints	# Avg. Violation (Å)	# Max. Violation (Å)
dpca01	263,919	36 (34.3%)	3.475	7.024
dpca02	170,517	39 (37.1%)	6.805	19.959
dpca03	130,588	41 (39.0%)	2.518	8.476
dpca04	120,401	37 (35.2%)	3.599	7.216
dpca05	133,577	41 (39.0%)	2.850	9.841
dpca06	123,575	40 (38.1%)	5.356	16.938
dpca07	159,039	33 (31.4%)	2.962	7.090
dpca08	73,161	37 (35.2%)	4.134	9.418
dpca09	106,973	36 (34.3%)	2.286	6.709
dpca10	37,756	41 (39.0%)	2.174	6.830
Full Simulation (results_2n0x)		26 (24.8%)	1.182	4.081

Table 6.1: Presentation of divergences derived from the results of each cluster's analysis and that of the full simulation, concerning differences in violated restraints and their respective violations

The Table presented above displays gcrarna runs from each individual cluster's analysis derived by comparing different frame data sizes for 2n0x simulations. The results strongly indicate that the Full Simulation has an overall substantially more accurate presentation, displaying fewer violated restraints, a lower violation percentage and much smaller average and maximum violations. On average the other runs have about a 46% difference of violated restraints on average compared to the full results. Particularly the average violation magnitude is worse: 3.616 Å vs. 1.182 Å, signaling that DPCAs' average violation is almost 3.06 times larger. The worst case violation is 2.44 times larger (9.95 Å vs. 4.081 Å) suggesting that overall the full simulation is more constraint-compliant and produces significantly less geometric variation.

Comparison of Representative MD Clusters with the Experimental NMR Structure



Figure 6.1 Structural superposition of the representative dPCA clusters onto the experimental solution NMR ensemble (PDB: 2N0X). The backbone atoms of the 20 NMR models are rendered as thin gray lines with CPK elemental coloring. The representative MD structures are highlighted in green for cluster 10 and cyan for cluster 09, corresponding to the clusters with the lowest restraint violations. The overlay demonstrates strong conservation along the central structural core, with conformational variability localized primarily at the flexible peptide termini.

For an assessment of whether or not the conformational ensembles sampled during molecular dynamics simulations agree with the experimental data, a structural superposition was performed between the representative structures of the clusters that appeared to display the lowest NOE violations (dPCA cluster 09 and cluster 10) and the experimental NMR structure from PDB (2N0X.pdb).

As shown in **Figure 6.1**, the representative clusters capture the overall fold and backbone topology of the experimental structure reasonably well.

- **Core Backbone Fold:** Both representative models adopt the characteristic U-turn loop geometry defined by the experimental NMR bundle, with their backbones embedded directly within the conformational envelope sampled across the 20 NMR models. This confirms that the clusters characterized by minimal NOE restraint violations reflect stable, experimentally supported conformational states rather than simulation artifacts.
- **Local Deviations and Terminal Flexibility:** Noticeable conformational divergence between the two clusters, and relative to the NMR bundle, is primarily concentrated at the unconstrained N- and C-termini. For instance, cluster 09 (cyan) displays a more extended conformation toward one terminus, whereas cluster 10 (green) remains noticeably more compact. This deviation is entirely consistent with the increased thermal mobility and conformational freedom characteristic of terminal residues in solution, which is likewise mirrored by the structural fanning of the NMR ensemble in those same segments.

Backbone c_{α} alignment yielded a root-mean-square deviation of $RMSD = 2.85 \text{ \AA}$ (following outlier rejection of the flexible terminal positions). Structural convergence is notably pronounced along the central structural core, whereas the primary deviations are localized at the flexible N- and C-termini as is above stated. This behavior is in line with the heightened thermal fluctuations and intrinsic conformational freedom typical of terminal residues in solution. Consequently, this alignment demonstrates that the MD clusters used with minimal restraint violations reflect physically realistic conformations compatible with the experimental NMR constraints rather than computational artifacts.

Core Divergence Insights

1. The Ensemble Averaging Effect (r^{-6} Weighting)

The data displays a phenomenon where the **Full Simulation** yields significantly fewer total violations (24.8%) and a lower maximum violation (4.081 Å) than any single isolated cluster. That is particularly encouraging and not necessarily indicative of a more rigid or homogenous structure. Instead it's consistent with the time averaging inherent in NOE measurements, where experimental NOEs present on the ensemble averaged distances sampled by the peptide over different time checkpoints of the experiment. In a flexible peptide different conformations can therefore contribute to the satisfaction of a given NOE restraint, with the short-distance conformations having a particularly strong contribution under the r^{-6} distance dependence. The fact that the best agreement with the experimental restraints is obtained for the complete trajectory is consequently an encouraging result: it indicates that the simulation samples the structural heterogeneity of the peptide in a manner that is consistent with the experimentally observed ensemble. By contrast, when the trajectory is partitioned into individual dPCA clusters using grcarma [Koukos & Glykos, 2013], this ensemble averaging is substantially reduced. Individual clusters preferentially represent particular conformational substates, including more open or alternative conformations in which some of the short-range contacts required by the NOEs are less frequently sampled. Their apparent increase in restraint violations therefore reflects, at least in part, the loss of conformational averaging rather than a failure of the overall simulation to reproduce the experimental data.

For a flexible peptide, better agreement of the complete trajectory than of any individual cluster is therefore a positive result: it suggests that the simulation reproduces not merely a particular structure, but the conformational ensemble sampled by the peptide.

2. Extreme Outliers in Specific Clusters (DPCA02 & DPCA06)

- The **DPCA02** and **DPCA06** clusters represent heavily distorted or nonnative conformational subsets.
- **DPCA02** also exhibits a serious max violation of **19.959 Å** involving the terminal hydrophobic pairing (1 LEU QD2 $\rightleftharpoons$ 9 PRO HB2).
- A similar topology is observed with **DPCA06**, with a maximum violation of **16.938 Å** between exactly the same two atoms.

3. Population and Resolution Trends

- In addition, **DPCA01** still shows high violations of 34.3% despite having the largest amount of frames (263,919) and a max deviation of 7.024 Å. These results prove that even the most populated structural state in the simulation is not sufficient enough to reproduce the spatial averages of the experimental values.
- Even the tightest cluster by average violation scale (**DPCA10** at 2.174 Å demonstrates a 39.0% total restraint violation rate, illustrating that structural refinement cannot rely on a single dominant dPCA conformer cluster to explain the experimental solution state.

Once again as with the previous case, the tabular results of each analysis are formatted and presented below for clear identification:

• **2NOX Full Simulation:**

NOE Analysis Results:

EXP	<r ⁻³ >	<r ⁻⁶ >	Deviation	Violation	Atoms Involved			
5.14	5.180	4.511	-0.629	0.000	10 GLN	HB2 <==>	12 SER	H
5.14	4.721	3.949	-1.191	0.000	10 GLN	HB3 <==>	12 SER	H
5.41	3.237	2.956	-2.454	0.000	11 VAL	HB <==>	12 SER	H
5.50	2.956	2.768	-2.732	0.000	10 GLN	H <==>	10 GLN	QG
5.50	5.294	4.128	-1.372	0.000	10 GLN	QG <==>	12 SER	H
5.50	7.614	5.486	-0.014	0.000	5 THR	H <==>	13 THR	QG2
5.50	3.083	2.864	-2.636	0.000	13 THR	H <==>	13 THR	QG2
4.92	2.963	2.810	-2.110	0.000	15 THR	H <==>	15 THR	QG2
5.50	2.924	2.726	-2.774	0.000	7 LYS	H <==>	7 LYS	QG
5.50	3.938	3.583	-1.917	0.000	6 LYS	QG <==>	7 LYS	H
5.50	4.359	4.045	-1.455	0.000	7 LYS	H <==>	7 LYS	QD
5.50	2.874	2.708	-2.792	0.000	6 LYS	H <==>	6 LYS	QG
5.50	4.285	3.921	-1.579	0.000	6 LYS	H <==>	6 LYS	QD
5.50	5.048	4.775	-0.725	0.000	6 LYS	H <==>	6 LYS	QE
3.75	4.610	4.143	0.393	0.393	7 LYS	HA <==>	7 LYS	QE
3.08	4.573	4.063	0.983	0.983	6 LYS	HA <==>	6 LYS	QE
2.97	2.455	2.443	-0.527	0.000	6 LYS	HB2 <==>	6 LYS	QG
2.88	2.459	2.446	-0.434	0.000	7 LYS	HB2 <==>	7 LYS	QG
5.50	3.271	2.955	-2.545	0.000	3 ARG	H <==>	3 ARG	HG2
5.50	3.211	2.982	-2.518	0.000	3 ARG	H <==>	3 ARG	HG3
4.33	3.547	3.063	-1.267	0.000	3 ARG	HA <==>	3 ARG	QD
3.86	3.197	2.958	-0.902	0.000	7 LYS	HB3 <==>	8 VAL	H
4.08	2.421	2.343	-1.737	0.000	8 VAL	H <==>	8 VAL	QG2
5.15	3.257	3.021	-2.129	0.000	10 GLN	HB3 <==>	11 VAL	H
3.11	2.365	2.306	-0.804	0.000	9 PRO	HA <==>	10 GLN	H
3.18	2.429	2.360	-0.820	0.000	14 PRO	HA <==>	15 THR	H
3.36	2.764	2.633	-0.727	0.000	12 SER	HA <==>	13 THR	H
5.50	4.917	4.620	-0.880	0.000	8 VAL	H <==>	9 PRO	HD2
5.50	2.833	2.681	-2.819	0.000	5 THR	H <==>	5 THR	QG2
2.92	2.552	2.440	-0.480	0.000	15 THR	HA <==>	16 LEU	H

2.92	2.634	2.597	-0.323	0.000	13 THR	HA <==>	13 THR	HB
5.50	2.880	2.705	-2.795	0.000	2 VAL	HB <==>	3 ARG	H
5.50	3.220	2.827	-2.673	0.000	4 TYR	H <==>	5 THR	H
2.86	3.234	3.043	0.183	0.183	8 VAL	H <==>	8 VAL	HB
4.28	2.915	2.601	-1.679	0.000	1 LEU	HA <==>	1 LEU	QD1
4.28	2.590	2.405	-1.875	0.000	1 LEU	HA <==>	1 LEU	QD2
4.65	6.141	4.437	-0.213	0.000	1 LEU	QD1 <==>	4 TYR	QD
4.65	6.707	4.797	0.147	0.147	1 LEU	QD2 <==>	4 TYR	QD
5.50	6.317	4.495	-1.005	0.000	1 LEU	QD1 <==>	4 TYR	QE
5.50	7.026	5.000	-0.500	0.000	1 LEU	QD2 <==>	4 TYR	QE
4.77	5.283	4.467	-0.303	0.000	3 ARG	HB3 <==>	4 TYR	QE
3.65	2.990	2.870	-0.780	0.000	3 ARG	HA <==>	3 ARG	HG3
3.65	2.867	2.754	-0.896	0.000	3 ARG	HA <==>	3 ARG	HG2
4.77	5.602	4.676	-0.094	0.000	3 ARG	HB2 <==>	4 TYR	QE
5.29	4.511	4.045	-1.245	0.000	3 ARG	HA <==>	4 TYR	QD
3.17	2.727	2.627	-0.543	0.000	4 TYR	HA <==>	4 TYR	QD
5.11	5.744	4.415	-0.695	0.000	4 TYR	QD <==>	9 PRO	HD3
5.50	5.353	3.916	-1.584	0.000	4 TYR	QE <==>	9 PRO	HD3
2.84	2.526	2.484	-0.356	0.000	13 THR	HA <==>	13 THR	QG2
2.40	2.370	2.363	-0.037	0.000	7 LYS	QG <==>	7 LYS	QD
2.88	2.459	2.447	-0.433	0.000	7 LYS	HB3 <==>	7 LYS	QG
3.82	2.438	2.409	-1.411	0.000	7 LYS	QG <==>	7 LYS	QE
3.79	3.484	3.038	-0.752	0.000	7 LYS	HA <==>	7 LYS	QD
3.63	2.636	2.570	-1.060	0.000	7 LYS	HA <==>	7 LYS	QG
2.97	2.460	2.448	-0.522	0.000	6 LYS	HB3 <==>	6 LYS	QG
2.40	2.370	2.363	-0.037	0.000	6 LYS	QG <==>	6 LYS	QD
3.63	3.451	3.015	-0.615	0.000	6 LYS	HA <==>	6 LYS	QD
3.45	2.649	2.579	-0.871	0.000	6 LYS	HA <==>	6 LYS	QG
5.50	5.691	5.545	0.045	0.045	7 LYS	HA <==>	8 VAL	HB
3.86	2.703	2.624	-1.236	0.000	10 GLN	HA <==>	10 GLN	QG
3.68	3.267	3.073	-0.607	0.000	11 VAL	H <==>	11 VAL	HB
4.56	3.288	3.031	-1.529	0.000	12 SER	QB <==>	13 THR	H
5.50	5.051	4.600	-0.900	0.000	7 LYS	QD <==>	8 VAL	H
3.86	3.558	3.316	-0.544	0.000	7 LYS	HB2 <==>	8 VAL	H
2.85	2.636	2.502	-0.348	0.000	7 LYS	HA <==>	8 VAL	H
5.50	4.733	4.434	-1.066	0.000	14 PRO	HA <==>	16 LEU	H
2.40	2.267	2.229	-0.171	0.000	8 VAL	HA <==>	9 PRO	HD3
2.41	2.573	2.506	0.096	0.096	8 VAL	HA <==>	9 PRO	HD2
3.46	10.088	7.541	4.081	4.081	5 THR	HA <==>	14 PRO	HD2
3.46	10.209	7.260	3.800	3.800	5 THR	HA <==>	14 PRO	HD3
4.23	8.187	5.476	1.246	1.246	5 THR	QG2 <==>	14 PRO	HD2
4.11	2.665	2.502	-1.608	0.000	8 VAL	HB <==>	9 PRO	HD2
5.01	9.375	6.540	1.530	1.530	1 LEU	QD1 <==>	9 PRO	HA
5.01	9.706	6.929	1.919	1.919	1 LEU	QD2 <==>	9 PRO	HA
5.50	10.460	6.787	1.287	1.287	1 LEU	QD2 <==>	9 PRO	HD3
5.50	9.971	6.334	0.834	0.834	1 LEU	QD1 <==>	9 PRO	HD3
4.47	8.174	5.238	0.768	0.768	1 LEU	QD1 <==>	9 PRO	HB3
4.47	8.670	5.532	1.062	1.062	1 LEU	QD2 <==>	9 PRO	HB3
4.47	7.731	4.840	0.370	0.370	1 LEU	QD1 <==>	9 PRO	HB2
4.47	8.491	5.274	0.804	0.804	1 LEU	QD2 <==>	9 PRO	HB2
4.08	3.082	2.861	-1.219	0.000	8 VAL	H <==>	8 VAL	QG1
3.23	3.130	2.945	-0.285	0.000	8 VAL	HA <==>	8 VAL	QG2
3.23	2.471	2.442	-0.788	0.000	8 VAL	HA <==>	8 VAL	QG1
5.15	3.511	3.289	-1.861	0.000	10 GLN	HB2 <==>	11 VAL	H
5.50	7.392	5.792	0.292	0.292	4 TYR	QD <==>	9 PRO	HG2
5.14	7.229	5.598	0.458	0.458	4 TYR	QE <==>	9 PRO	HG2
5.20	5.851	4.156	-1.044	0.000	4 TYR	QD <==>	13 THR	QG2
4.23	8.649	5.835	1.605	1.605	5 THR	QG2 <==>	14 PRO	HD3
3.05	6.690	4.503	1.453	1.453	5 THR	QG2 <==>	13 THR	QG2
5.34	7.062	6.722	1.382	1.382	1 LEU	HA <==>	3 ARG	QB
2.24	2.375	2.368	0.128	0.128	3 ARG	QB <==>	3 ARG	QG
3.03	2.477	2.451	-0.579	0.000	3 ARG	QB <==>	3 ARG	QD
4.97	4.229	3.663	-1.307	0.000	3 ARG	QB <==>	4 TYR	QD

5.34	4.840	4.072	-1.268	0.000	3 ARG	QG <==>	4 TYR	QD
5.34	5.392	4.069	-1.271	0.000	4 TYR	QD <==>	9 PRO	QB
4.33	9.093	7.100	2.770	2.770	5 THR	H <==>	14 PRO	QD
3.76	9.639	6.660	2.900	2.900	5 THR	HB <==>	14 PRO	QD
5.35	7.718	5.058	-0.292	0.000	5 THR	QG2 <==>	14 PRO	QG
2.73	2.479	2.458	-0.272	0.000	6 LYS	QB <==>	6 LYS	QD
3.64	3.108	2.732	-0.908	0.000	6 LYS	QB <==>	6 LYS	QE
2.74	2.477	2.456	-0.284	0.000	7 LYS	QB <==>	7 LYS	QD
2.53	3.112	2.736	0.206	0.206	7 LYS	QB <==>	7 LYS	QE
5.32	4.604	4.558	-0.762	0.000	9 PRO	QB <==>	10 GLN	HA
2.63	2.457	2.445	-0.185	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.604	4.568	-0.772	0.000	14 PRO	QB <==>	15 THR	HA

=====

SUMMARY: 105 unique NOE restraints displayed.
Violations (>0.001Å): 26 (24.8%)
Average violation (for violated restraints): 1.182 Å
Maximum violation: 4.081 Å
Violation severity distribution:
0-0.5Å: 10 restraints
>0.5Å: 4 restraints
>1.0Å: 8 restraints
>2.0Å: 2 restraints
>3.0Å: 2 restraints

● *DPCA CLUSTERING ANALYSIS (OF THE FIRST 10 CLUSTERS) ON 2N0X:*

1. *DPCA01 (FIRST CLUSTER):*

Starting NOE processing using 'carma.distances' at 2025-11-13 01:53:20
Defined 105 unique restraint groups from 257 calculated distances.
Processing 263,919 frames from 'carma.distances' in memory...
Progress: 263919/263919 (100.0%) | Elapsed: 23.5s | Remaining: 0.0s
Finished processing 263,919 frames.

NOE Analysis Results:

=====

EXP	<r^-3>	<r^-6>	Deviation	Violation	Atoms Involved
-----	--------	--------	-----------	-----------	----------------

=====									
5.14	5.854	5.814	0.674	0.674	10 GLN	HB2 <==>	12 SER	H	
5.14	6.127	6.080	0.940	0.940	10 GLN	HB3 <==>	12 SER	H	
5.41	2.744	2.644	-2.766	0.000	11 VAL	HB <==>	12 SER	H	
5.50	3.128	3.052	-2.448	0.000	10 GLN	H <==>	10 GLN	QG	
5.50	6.636	6.594	1.094	1.094	10 GLN	QG <==>	12 SER	H	
5.50	5.829	5.800	0.300	0.300	5 THR	H <==>	13 THR	QG2	
5.50	2.671	2.630	-2.870	0.000	13 THR	H <==>	13 THR	QG2	
4.92	2.916	2.805	-2.115	0.000	15 THR	H <==>	15 THR	QG2	
5.50	3.060	2.919	-2.581	0.000	7 LYS	H <==>	7 LYS	QG	
5.50	4.794	4.722	-0.778	0.000	6 LYS	QG <==>	7 LYS	H	
5.50	4.359	4.128	-1.372	0.000	7 LYS	H <==>	7 LYS	QD	
5.50	2.939	2.859	-2.641	0.000	6 LYS	H <==>	6 LYS	QG	
5.50	4.249	3.714	-1.786	0.000	6 LYS	H <==>	6 LYS	QD	
5.50	5.087	4.859	-0.641	0.000	6 LYS	H <==>	6 LYS	QE	
3.75	4.618	4.255	0.505	0.505	7 LYS	HA <==>	7 LYS	QE	
3.08	4.597	4.176	1.096	1.096	6 LYS	HA <==>	6 LYS	QE	
2.97	2.458	2.446	-0.524	0.000	6 LYS	HB2 <==>	6 LYS	QG	
2.88	2.445	2.432	-0.448	0.000	7 LYS	HB2 <==>	7 LYS	QG	
5.50	4.119	3.803	-1.697	0.000	3 ARG	H <==>	3 ARG	HG2	
5.50	3.740	3.311	-2.189	0.000	3 ARG	H <==>	3 ARG	HG3	
4.33	3.415	2.961	-1.369	0.000	3 ARG	HA <==>	3 ARG	QD	
3.86	2.494	2.448	-1.412	0.000	7 LYS	HB3 <==>	8 VAL	H	
4.08	2.399	2.349	-1.731	0.000	8 VAL	H <==>	8 VAL	QG2	
5.15	3.949	3.758	-1.392	0.000	10 GLN	HB3 <==>	11 VAL	H	
3.11	2.134	2.121	-0.989	0.000	9 PRO	HA <==>	10 GLN	H	
3.18	2.327	2.307	-0.873	0.000	14 PRO	HA <==>	15 THR	H	
3.36	3.343	3.332	-0.028	0.000	12 SER	HA <==>	13 THR	H	
5.50	5.143	5.126	-0.374	0.000	8 VAL	H <==>	9 PRO	HD2	
5.50	2.567	2.497	-3.003	0.000	5 THR	H <==>	5 THR	QG2	
2.92	2.273	2.257	-0.663	0.000	15 THR	HA <==>	16 LEU	H	
2.92	2.383	2.374	-0.546	0.000	13 THR	HA <==>	13 THR	HB	
5.50	2.617	2.573	-2.927	0.000	2 VAL	HB <==>	3 ARG	H	
5.50	4.426	4.419	-1.081	0.000	4 TYR	H <==>	5 THR	H	
2.86	2.560	2.531	-0.329	0.000	8 VAL	H <==>	8 VAL	HB	
4.28	3.157	2.748	-1.532	0.000	1 LEU	HA <==>	1 LEU	QD1	
4.28	2.452	2.327	-1.953	0.000	1 LEU	HA <==>	1 LEU	QD2	
4.65	11.465	11.231	6.581	6.581	1 LEU	QD1 <==>	4 TYR	QD	
4.65	11.393	11.220	6.570	6.570	1 LEU	QD2 <==>	4 TYR	QD	
5.50	12.720	12.518	7.018	7.018	1 LEU	QD1 <==>	4 TYR	QE	
5.50	12.745	12.524	7.024	7.024	1 LEU	QD2 <==>	4 TYR	QE	
4.77	3.877	3.732	-1.038	0.000	3 ARG	HB3 <==>	4 TYR	QE	
3.65	2.611	2.550	-1.100	0.000	3 ARG	HA <==>	3 ARG	HG3	
3.65	2.983	2.889	-0.761	0.000	3 ARG	HA <==>	3 ARG	HG2	
4.77	4.980	4.667	-0.103	0.000	3 ARG	HB2 <==>	4 TYR	QE	
5.29	5.324	5.270	-0.020	0.000	3 ARG	HA <==>	4 TYR	QD	
3.17	2.641	2.578	-0.592	0.000	4 TYR	HA <==>	4 TYR	QD	
5.11	9.150	8.280	3.170	3.170	4 TYR	QD <==>	9 PRO	HD3	
5.50	11.024	9.933	4.433	4.433	4 TYR	QE <==>	9 PRO	HD3	
2.84	3.691	3.637	0.797	0.797	13 THR	HA <==>	13 THR	QG2	
2.40	2.376	2.370	-0.030	0.000	7 LYS	QG <==>	7 LYS	QD	
2.88	2.457	2.445	-0.435	0.000	7 LYS	HB3 <==>	7 LYS	QG	
3.82	2.395	2.373	-1.447	0.000	7 LYS	QG <==>	7 LYS	QE	
3.79	3.269	2.883	-0.907	0.000	7 LYS	HA <==>	7 LYS	QD	
3.63	2.527	2.483	-1.147	0.000	7 LYS	HA <==>	7 LYS	QG	
2.97	2.462	2.450	-0.520	0.000	6 LYS	HB3 <==>	6 LYS	QG	
2.40	2.371	2.364	-0.036	0.000	6 LYS	QG <==>	6 LYS	QD	
3.63	3.515	3.078	-0.552	0.000	6 LYS	HA <==>	6 LYS	QD	
3.45	2.586	2.537	-0.913	0.000	6 LYS	HA <==>	6 LYS	QG	
5.50	4.745	4.717	-0.783	0.000	7 LYS	HA <==>	8 VAL	HB	
3.86	2.776	2.691	-1.169	0.000	10 GLN	HA <==>	10 GLN	QG	
3.68	2.681	2.602	-1.078	0.000	11 VAL	H <==>	11 VAL	HB	
4.56	4.073	3.959	-0.601	0.000	12 SER	QB <==>	13 THR	H	

5.50	4.428	4.294	-1.206	0.000	7 LYS	QD <==>	8 VAL	H
3.86	3.753	3.718	-0.142	0.000	7 LYS	HB2 <==>	8 VAL	H
2.85	2.372	2.353	-0.497	0.000	7 LYS	HA <==>	8 VAL	H
5.50	6.404	6.396	0.896	0.896	14 PRO	HA <==>	16 LEU	H
2.40	2.238	2.202	-0.198	0.000	8 VAL	HA <==>	9 PRO	HD3
2.41	2.421	2.372	-0.038	0.000	8 VAL	HA <==>	9 PRO	HD2
3.46	8.526	8.436	4.976	4.976	5 THR	HA <==>	14 PRO	HD2
3.46	9.737	9.649	6.189	6.189	5 THR	HA <==>	14 PRO	HD3
4.23	9.551	9.430	5.200	5.200	5 THR	QG2 <==>	14 PRO	HD2
4.11	4.106	3.673	-0.437	0.000	8 VAL	HB <==>	9 PRO	HD2
5.01	10.513	9.698	4.688	4.688	1 LEU	QD1 <==>	9 PRO	HA
5.01	9.890	9.079	4.069	4.069	1 LEU	QD2 <==>	9 PRO	HA
5.50	9.140	7.463	1.963	1.963	1 LEU	QD2 <==>	9 PRO	HD3
5.50	10.302	8.941	3.441	3.441	1 LEU	QD1 <==>	9 PRO	HD3
4.47	8.863	7.071	2.601	2.601	1 LEU	QD1 <==>	9 PRO	HB3
4.47	8.103	6.228	1.758	1.758	1 LEU	QD2 <==>	9 PRO	HB3
4.47	10.103	8.049	3.579	3.579	1 LEU	QD1 <==>	9 PRO	HB2
4.47	9.348	7.369	2.899	2.899	1 LEU	QD2 <==>	9 PRO	HB2
4.08	4.213	3.911	-0.169	0.000	8 VAL	H <==>	8 VAL	QG1
3.23	2.551	2.518	-0.712	0.000	8 VAL	HA <==>	8 VAL	QG2
3.23	2.548	2.525	-0.705	0.000	8 VAL	HA <==>	8 VAL	QG1
5.15	3.511	3.367	-1.783	0.000	10 GLN	HB2 <==>	11 VAL	H
5.50	10.890	10.213	4.713	4.713	4 TYR	QD <==>	9 PRO	HG2
5.14	12.439	9.637	4.497	4.497	4 TYR	QE <==>	9 PRO	HG2
5.20	4.839	4.492	-0.708	0.000	4 TYR	QD <==>	13 THR	QG2
4.23	10.978	10.876	6.646	6.646	5 THR	QG2 <==>	14 PRO	HD3
3.05	6.785	6.651	3.601	3.601	5 THR	QG2 <==>	13 THR	QG2
5.34	8.159	8.126	2.786	2.786	1 LEU	HA <==>	3 ARG	QB
2.24	2.380	2.374	0.134	0.134	3 ARG	QB <==>	3 ARG	QG
3.03	2.511	2.486	-0.544	0.000	3 ARG	QB <==>	3 ARG	QD
4.97	3.405	3.261	-1.709	0.000	3 ARG	QB <==>	4 TYR	QD
5.34	3.705	3.473	-1.867	0.000	3 ARG	QG <==>	4 TYR	QD
5.34	9.285	9.062	3.722	3.722	4 TYR	QD <==>	9 PRO	QB
4.33	7.854	7.798	3.468	3.468	5 THR	H <==>	14 PRO	QD
3.76	10.558	10.486	6.726	6.726	5 THR	HB <==>	14 PRO	QD
5.35	10.974	10.817	5.467	5.467	5 THR	QG2 <==>	14 PRO	QG
2.73	2.472	2.452	-0.278	0.000	6 LYS	QB <==>	6 LYS	QD
3.64	3.078	2.711	-0.929	0.000	6 LYS	QB <==>	6 LYS	QE
2.74	2.486	2.469	-0.271	0.000	7 LYS	QB <==>	7 LYS	QD
2.53	3.826	3.390	0.860	0.860	7 LYS	QB <==>	7 LYS	QE
5.32	5.086	5.067	-0.253	0.000	9 PRO	QB <==>	10 GLN	HA
2.63	2.436	2.424	-0.206	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.620	4.603	-0.737	0.000	14 PRO	QB <==>	15 THR	HA

```

=====
SUMMARY: 105 unique NOE restraints displayed.
Violations (>0.001Å): 36 (34.3%)
Average violation (for violated restraints): 3.475 Å
Maximum violation: 7.024 Å
Violation severity distribution:
  0-0.5Å: 2 restraints
  >0.5Å: 6 restraints
  >1.0Å: 4 restraints
  >2.0Å: 3 restraints
  >3.0Å: 21 restraints

```

2. DPCA02 (SECOND CLUSTER):

Starting NOE processing using 'carma.distances' at 2025-11-13 01:52:04

Defined 105 unique restraint groups from 257 calculated distances.
Processing 170,517 frames from 'carma.distances' in memory...
Progress: 170517/170517 (100.0%) | Elapsed: 15.5s | Remaining: 0.0s
Finished processing 170,517 frames.

NOE Analysis Results:

EXP	<r ⁻³ >	<r ⁻⁶ >	Deviation	Violation	Atoms Involved				
5.14	5.376	5.337	0.197	0.197	10	GLN	HB2 <==>	12	SER H
5.14	5.552	5.505	0.365	0.365	10	GLN	HB3 <==>	12	SER H
5.41	4.055	3.961	-1.449	0.000	11	VAL	HB <==>	12	SER H
5.50	3.183	3.072	-2.428	0.000	10	GLN	H <==>	10	GLN QG
5.50	6.150	6.108	0.608	0.608	10	GLN	QG <==>	12	SER H
5.50	9.660	9.512	4.012	4.012	5	THR	H <==>	13	THR QG2
5.50	3.538	3.176	-2.324	0.000	13	THR	H <==>	13	THR QG2
4.92	3.144	3.023	-1.897	0.000	15	THR	H <==>	15	THR QG2
5.50	3.110	2.998	-2.502	0.000	7	LYS	H <==>	7	LYS QG
5.50	3.610	3.346	-2.154	0.000	6	LYS	QG <==>	7	LYS H
5.50	4.627	4.294	-1.206	0.000	7	LYS	H <==>	7	LYS QD
5.50	2.993	2.870	-2.630	0.000	6	LYS	H <==>	6	LYS QG
5.50	3.846	3.254	-2.246	0.000	6	LYS	H <==>	6	LYS QD
5.50	4.623	4.304	-1.196	0.000	6	LYS	H <==>	6	LYS QE
3.75	4.978	4.907	1.157	1.157	7	LYS	HA <==>	7	LYS QE
3.08	4.951	4.744	1.664	1.664	6	LYS	HA <==>	6	LYS QE
2.97	2.461	2.447	-0.523	0.000	6	LYS	HB2 <==>	6	LYS QG
2.88	2.481	2.470	-0.410	0.000	7	LYS	HB2 <==>	7	LYS QG
5.50	3.331	2.954	-2.546	0.000	3	ARG	H <==>	3	ARG HG2
5.50	3.529	3.233	-2.267	0.000	3	ARG	H <==>	3	ARG HG3
4.33	3.223	2.809	-1.521	0.000	3	ARG	HA <==>	3	ARG QD
3.86	2.998	2.886	-0.974	0.000	7	LYS	HB3 <==>	8	VAL H
4.08	2.304	2.265	-1.815	0.000	8	VAL	H <==>	8	VAL QG2
5.15	3.550	3.362	-1.788	0.000	10	GLN	HB3 <==>	11	VAL H
3.11	2.238	2.223	-0.887	0.000	9	PRO	HA <==>	10	GLN H
3.18	2.225	2.213	-0.967	0.000	14	PRO	HA <==>	15	THR H
3.36	3.086	3.065	-0.295	0.000	12	SER	HA <==>	13	THR H
5.50	5.175	5.158	-0.342	0.000	8	VAL	H <==>	9	PRO HD2
5.50	3.164	3.095	-2.405	0.000	5	THR	H <==>	5	THR QG2
2.92	2.321	2.300	-0.620	0.000	15	THR	HA <==>	16	LEU H
2.92	2.866	2.826	-0.094	0.000	13	THR	HA <==>	13	THR HB
5.50	2.806	2.679	-2.821	0.000	2	VAL	HB <==>	3	ARG H
5.50	4.427	4.418	-1.082	0.000	4	TYR	H <==>	5	THR H
2.86	2.605	2.556	-0.304	0.000	8	VAL	H <==>	8	VAL HB
4.28	2.850	2.562	-1.718	0.000	1	LEU	HA <==>	1	LEU QD1
4.28	2.631	2.430	-1.850	0.000	1	LEU	HA <==>	1	LEU QD2
4.65	10.287	8.499	3.849	3.849	1	LEU	QD1 <==>	4	TYR QD
4.65	10.310	8.301	3.651	3.651	1	LEU	QD2 <==>	4	TYR QD
5.50	10.864	8.201	2.701	2.701	1	LEU	QD1 <==>	4	TYR QE
5.50	10.848	8.395	2.895	2.895	1	LEU	QD2 <==>	4	TYR QE
4.77	4.293	4.006	-0.764	0.000	3	ARG	HB3 <==>	4	TYR QE
3.65	2.909	2.789	-0.861	0.000	3	ARG	HA <==>	3	ARG HG3
3.65	3.043	2.912	-0.738	0.000	3	ARG	HA <==>	3	ARG HG2
4.77	4.756	4.372	-0.398	0.000	3	ARG	HB2 <==>	4	TYR QE
5.29	5.335	5.277	-0.013	0.000	3	ARG	HA <==>	4	TYR QD
3.17	2.711	2.598	-0.572	0.000	4	TYR	HA <==>	4	TYR QD
5.11	11.532	11.399	6.289	6.289	4	TYR	QD <==>	9	PRO HD3
5.50	12.441	12.143	6.643	6.643	4	TYR	QE <==>	9	PRO HD3
2.84	2.482	2.459	-0.381	0.000	13	THR	HA <==>	13	THR QG2
2.40	2.371	2.364	-0.036	0.000	7	LYS	QG <==>	7	LYS QD
2.88	2.460	2.449	-0.431	0.000	7	LYS	HB3 <==>	7	LYS QG
3.82	2.425	2.400	-1.420	0.000	7	LYS	QG <==>	7	LYS QE
3.79	4.318	4.116	0.326	0.326	7	LYS	HA <==>	7	LYS QD
3.63	2.640	2.576	-1.054	0.000	7	LYS	HA <==>	7	LYS QG

2.97	2.456	2.444	-0.526	0.000	6 LYS	HB3 <==>	6 LYS	QG
2.40	2.369	2.363	-0.037	0.000	6 LYS	QG <==>	6 LYS	QD
3.63	4.122	3.708	0.078	0.078	6 LYS	HA <==>	6 LYS	QD
3.45	2.767	2.675	-0.775	0.000	6 LYS	HA <==>	6 LYS	QG
5.50	4.786	4.745	-0.755	0.000	7 LYS	HA <==>	8 VAL	HB
3.86	2.604	2.551	-1.309	0.000	10 GLN	HA <==>	10 GLN	QG
3.68	3.581	3.458	-0.222	0.000	11 VAL	H <==>	11 VAL	HB
4.56	4.021	3.991	-0.569	0.000	12 SER	QB <==>	13 THR	H
5.50	5.065	4.992	-0.508	0.000	7 LYS	QD <==>	8 VAL	H
3.86	3.504	3.262	-0.598	0.000	7 LYS	HB2 <==>	8 VAL	H
2.85	2.218	2.205	-0.645	0.000	7 LYS	HA <==>	8 VAL	H
5.50	6.409	6.402	0.902	0.902	14 PRO	HA <==>	16 LEU	H
2.40	2.236	2.203	-0.197	0.000	8 VAL	HA <==>	9 PRO	HD3
2.41	2.489	2.436	0.026	0.026	8 VAL	HA <==>	9 PRO	HD2
3.46	10.613	10.568	7.108	7.108	5 THR	HA <==>	14 PRO	HD2
3.46	10.267	10.211	6.751	6.751	5 THR	HA <==>	14 PRO	HD3
4.23	11.813	11.775	7.545	7.545	5 THR	QG2 <==>	14 PRO	HD2
4.11	3.618	3.137	-0.973	0.000	8 VAL	HB <==>	9 PRO	HD2
5.01	22.858	22.186	17.176	17.176	1 LEU	QD1 <==>	9 PRO	HA
5.01	22.865	22.236	17.226	17.226	1 LEU	QD2 <==>	9 PRO	HA
5.50	21.487	21.081	15.581	15.581	1 LEU	QD2 <==>	9 PRO	HD3
5.50	21.444	21.001	15.501	15.501	1 LEU	QD1 <==>	9 PRO	HD3
4.47	24.145	23.568	19.098	19.098	1 LEU	QD1 <==>	9 PRO	HB3
4.47	24.157	23.652	19.182	19.182	1 LEU	QD2 <==>	9 PRO	HB3
4.47	24.841	24.374	19.904	19.904	1 LEU	QD1 <==>	9 PRO	HB2
4.47	24.855	24.429	19.959	19.959	1 LEU	QD2 <==>	9 PRO	HB2
4.08	3.705	3.249	-0.831	0.000	8 VAL	H <==>	8 VAL	QG1
3.23	2.624	2.573	-0.657	0.000	8 VAL	HA <==>	8 VAL	QG2
3.23	2.560	2.533	-0.697	0.000	8 VAL	HA <==>	8 VAL	QG1
5.15	3.278	3.146	-2.004	0.000	10 GLN	HB2 <==>	11 VAL	H
5.50	14.050	13.952	8.452	8.452	4 TYR	QD <==>	9 PRO	HG2
5.14	14.827	14.605	9.465	9.465	4 TYR	QE <==>	9 PRO	HG2
5.20	7.236	6.990	1.790	1.790	4 TYR	QD <==>	13 THR	QG2
4.23	11.854	11.818	7.588	7.588	5 THR	QG2 <==>	14 PRO	HD3
3.05	12.724	12.586	9.536	9.536	5 THR	QG2 <==>	13 THR	QG2
5.34	7.462	7.310	1.970	1.970	1 LEU	HA <==>	3 ARG	QB
2.24	2.376	2.369	0.129	0.129	3 ARG	QB <==>	3 ARG	QG
3.03	2.487	2.463	-0.567	0.000	3 ARG	QB <==>	3 ARG	QD
4.97	3.423	3.221	-1.749	0.000	3 ARG	QB <==>	4 TYR	QD
5.34	3.767	3.469	-1.871	0.000	3 ARG	QG <==>	4 TYR	QD
5.34	14.190	14.021	8.681	8.681	4 TYR	QD <==>	9 PRO	QB
4.33	8.734	8.696	4.366	4.366	5 THR	H <==>	14 PRO	QD
3.76	11.823	11.785	8.025	8.025	5 THR	HB <==>	14 PRO	QD
5.35	10.193	10.136	4.786	4.786	5 THR	QG2 <==>	14 PRO	QG
2.73	2.479	2.456	-0.274	0.000	6 LYS	QB <==>	6 LYS	QD
3.64	2.773	2.492	-1.148	0.000	6 LYS	QB <==>	6 LYS	QE
2.74	2.460	2.445	-0.295	0.000	7 LYS	QB <==>	7 LYS	QD
2.53	3.108	2.738	0.208	0.208	7 LYS	QB <==>	7 LYS	QE
5.32	4.764	4.740	-0.580	0.000	9 PRO	QB <==>	10 GLN	HA
2.63	2.444	2.432	-0.198	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.408	4.384	-0.956	0.000	14 PRO	QB <==>	15 THR	HA

=====

SUMMARY: 105 unique NOE restraints displayed.

Violations (>0.001Å): 39 (37.1%)

Average violation (for violated restraints): 6.805 Å

Maximum violation: 19.959 Å

Violation severity distribution:

0-0.5Å: 7 restraints

>0.5Å: 2 restraints

>1.0Å: 4 restraints

>2.0Å: 2 restraints

>3.0Å: 24 restraints

3. DPCA03 (THIRD CLUSTER):

Starting NOE processing using 'carma.distances' at 2025-11-13 01:54:26
Defined 105 unique restraint groups from 257 calculated distances.
Processing 130,588 frames from 'carma.distances' in memory...
Progress: 130588/130588 (100.0%) | Elapsed: 11.7s | Remaining: 0.0s
Finished processing 130,588 frames.

NOE Analysis Results:

EXP	<r ⁻³ >	<r ⁻⁶ >	Deviation	Violation	Atoms Involved				
5.14	6.444	6.381	1.241	1.241	10 GLN	HB2 <==>	12 SER	H	
5.14	6.005	5.934	0.794	0.794	10 GLN	HB3 <==>	12 SER	H	
5.41	2.562	2.492	-2.918	0.000	11 VAL	HB <==>	12 SER	H	
5.50	3.056	2.967	-2.533	0.000	10 GLN	H <==>	10 GLN	QG	
5.50	7.543	7.024	1.524	1.524	10 GLN	QG <==>	12 SER	H	
5.50	4.173	4.036	-1.464	0.000	5 THR	H <==>	13 THR	QG2	
5.50	2.609	2.542	-2.958	0.000	13 THR	H <==>	13 THR	QG2	
4.92	2.757	2.639	-2.281	0.000	15 THR	H <==>	15 THR	QG2	
5.50	2.926	2.722	-2.778	0.000	7 LYS	H <==>	7 LYS	QG	
5.50	4.040	3.930	-1.570	0.000	6 LYS	QG <==>	7 LYS	H	
5.50	4.286	4.062	-1.438	0.000	7 LYS	H <==>	7 LYS	QD	
5.50	2.908	2.837	-2.663	0.000	6 LYS	H <==>	6 LYS	QG	
5.50	4.430	4.325	-1.175	0.000	6 LYS	H <==>	6 LYS	QD	
5.50	5.217	5.104	-0.396	0.000	6 LYS	H <==>	6 LYS	QE	
3.75	4.478	3.960	0.210	0.210	7 LYS	HA <==>	7 LYS	QE	
3.08	4.662	4.168	1.088	1.088	6 LYS	HA <==>	6 LYS	QE	
2.97	2.459	2.449	-0.521	0.000	6 LYS	HB2 <==>	6 LYS	QG	
2.88	2.458	2.446	-0.434	0.000	7 LYS	HB2 <==>	7 LYS	QG	
5.50	2.943	2.708	-2.792	0.000	3 ARG	H <==>	3 ARG	HG2	
5.50	2.816	2.684	-2.816	0.000	3 ARG	H <==>	3 ARG	HG3	
4.33	3.907	3.372	-0.958	0.000	3 ARG	HA <==>	3 ARG	QD	
3.86	3.320	3.159	-0.701	0.000	7 LYS	HB3 <==>	8 VAL	H	
4.08	2.189	2.149	-1.931	0.000	8 VAL	H <==>	8 VAL	QG2	
5.15	2.681	2.602	-2.548	0.000	10 GLN	HB3 <==>	11 VAL	H	
3.11	2.238	2.224	-0.886	0.000	9 PRO	HA <==>	10 GLN	H	
3.18	2.340	2.322	-0.858	0.000	14 PRO	HA <==>	15 THR	H	
3.36	2.842	2.811	-0.549	0.000	12 SER	HA <==>	13 THR	H	
5.50	5.184	5.167	-0.333	0.000	8 VAL	H <==>	9 PRO	HD2	
5.50	3.157	3.092	-2.408	0.000	5 THR	H <==>	5 THR	QG2	
2.92	2.258	2.242	-0.678	0.000	15 THR	HA <==>	16 LEU	H	
2.92	2.505	2.491	-0.429	0.000	13 THR	HA <==>	13 THR	HB	
5.50	2.492	2.435	-3.065	0.000	2 VAL	HB <==>	3 ARG	H	
5.50	4.356	4.268	-1.232	0.000	4 TYR	H <==>	5 THR	H	
2.86	3.580	3.476	0.616	0.616	8 VAL	H <==>	8 VAL	HB	
4.28	3.484	2.988	-1.292	0.000	1 LEU	HA <==>	1 LEU	QD1	
4.28	2.354	2.270	-2.010	0.000	1 LEU	HA <==>	1 LEU	QD2	
4.65	7.621	6.750	2.100	2.100	1 LEU	QD1 <==>	4 TYR	QD	
4.65	7.352	6.819	2.169	2.169	1 LEU	QD2 <==>	4 TYR	QD	
5.50	6.794	5.560	0.060	0.060	1 LEU	QD1 <==>	4 TYR	QE	
5.50	6.277	5.684	0.184	0.184	1 LEU	QD2 <==>	4 TYR	QE	
4.77	7.226	6.802	2.032	2.032	3 ARG	HB3 <==>	4 TYR	QE	
3.65	3.141	3.037	-0.613	0.000	3 ARG	HA <==>	3 ARG	HG3	
3.65	2.839	2.730	-0.920	0.000	3 ARG	HA <==>	3 ARG	HG2	
4.77	7.181	6.718	1.948	1.948	3 ARG	HB2 <==>	4 TYR	QE	
5.29	3.970	3.777	-1.513	0.000	3 ARG	HA <==>	4 TYR	QD	
3.17	2.614	2.530	-0.640	0.000	4 TYR	HA <==>	4 TYR	QD	
5.11	7.977	7.286	2.176	2.176	4 TYR	QD <==>	9 PRO	HD3	
5.50	7.642	6.440	0.940	0.940	4 TYR	QE <==>	9 PRO	HD3	

2.84	2.459	2.439	-0.401	0.000	13 THR	HA <==>	13 THR	QG2
2.40	2.376	2.370	-0.030	0.000	7 LYS	QG <==>	7 LYS	QD
2.88	2.452	2.441	-0.439	0.000	7 LYS	HB3 <==>	7 LYS	QG
3.82	2.435	2.409	-1.411	0.000	7 LYS	QG <==>	7 LYS	QE
3.79	3.267	2.880	-0.910	0.000	7 LYS	HA <==>	7 LYS	QD
3.63	2.607	2.554	-1.076	0.000	7 LYS	HA <==>	7 LYS	QG
2.97	2.470	2.460	-0.510	0.000	6 LYS	HB3 <==>	6 LYS	QG
2.40	2.379	2.373	-0.027	0.000	6 LYS	QG <==>	6 LYS	QD
3.63	3.606	3.151	-0.479	0.000	6 LYS	HA <==>	6 LYS	QD
3.45	2.591	2.551	-0.899	0.000	6 LYS	HA <==>	6 LYS	QG
5.50	6.513	6.495	0.995	0.995	7 LYS	HA <==>	8 VAL	HB
3.86	2.601	2.567	-1.293	0.000	10 GLN	HA <==>	10 GLN	QG
3.68	3.595	3.439	-0.241	0.000	11 VAL	H <==>	11 VAL	HB
4.56	4.220	4.180	-0.380	0.000	12 SER	QB <==>	13 THR	H
5.50	5.371	5.286	-0.214	0.000	7 LYS	QD <==>	8 VAL	H
3.86	3.339	3.193	-0.667	0.000	7 LYS	HB2 <==>	8 VAL	H
2.85	3.421	3.389	0.539	0.539	7 LYS	HA <==>	8 VAL	H
5.50	6.292	6.168	0.668	0.668	14 PRO	HA <==>	16 LEU	H
2.40	2.233	2.203	-0.197	0.000	8 VAL	HA <==>	9 PRO	HD3
2.41	2.755	2.681	0.271	0.271	8 VAL	HA <==>	9 PRO	HD2
3.46	7.823	7.565	4.105	4.105	5 THR	HA <==>	14 PRO	HD2
3.46	8.957	8.651	5.191	5.191	5 THR	HA <==>	14 PRO	HD3
4.23	4.787	4.201	-0.029	0.000	5 THR	QG2 <==>	14 PRO	HD2
4.11	2.358	2.316	-1.794	0.000	8 VAL	HB <==>	9 PRO	HD2
5.01	13.917	12.837	7.827	7.827	1 LEU	QD1 <==>	9 PRO	HA
5.01	12.998	11.864	6.854	6.854	1 LEU	QD2 <==>	9 PRO	HA
5.50	15.527	13.976	8.476	8.476	1 LEU	QD2 <==>	9 PRO	HD3
5.50	16.403	13.767	8.267	8.267	1 LEU	QD1 <==>	9 PRO	HD3
4.47	14.292	10.967	6.497	6.497	1 LEU	QD1 <==>	9 PRO	HB3
4.47	13.198	10.888	6.418	6.418	1 LEU	QD2 <==>	9 PRO	HB3
4.47	15.320	9.984	5.514	5.514	1 LEU	QD1 <==>	9 PRO	HB2
4.47	14.361	11.847	7.377	7.377	1 LEU	QD2 <==>	9 PRO	HB2
4.08	2.622	2.549	-1.531	0.000	8 VAL	H <==>	8 VAL	QG1
3.23	3.628	3.516	0.286	0.286	8 VAL	HA <==>	8 VAL	QG2
3.23	2.437	2.417	-0.813	0.000	8 VAL	HA <==>	8 VAL	QG1
5.15	3.611	3.450	-1.700	0.000	10 GLN	HB2 <==>	11 VAL	H
5.50	8.715	8.317	2.817	2.817	4 TYR	QD <==>	9 PRO	HG2
5.14	7.952	7.334	2.194	2.194	4 TYR	QE <==>	9 PRO	HG2
5.20	3.145	2.912	-2.288	0.000	4 TYR	QD <==>	13 THR	QG2
4.23	5.788	5.120	0.890	0.890	5 THR	QG2 <==>	14 PRO	HD3
3.05	4.780	4.300	1.250	1.250	5 THR	QG2 <==>	13 THR	QG2
5.34	8.178	8.074	2.734	2.734	1 LEU	HA <==>	3 ARG	QB
2.24	2.388	2.382	0.142	0.142	3 ARG	QB <==>	3 ARG	QG
3.03	2.461	2.443	-0.587	0.000	3 ARG	QB <==>	3 ARG	QD
4.97	5.612	5.326	0.356	0.356	3 ARG	QB <==>	4 TYR	QD
5.34	6.222	5.788	0.448	0.448	3 ARG	QG <==>	4 TYR	QD
5.34	6.653	6.009	0.669	0.669	4 TYR	QD <==>	9 PRO	QB
4.33	6.366	6.028	1.698	1.698	5 THR	H <==>	14 PRO	QD
3.76	7.581	7.158	3.398	3.398	5 THR	HB <==>	14 PRO	QD
5.35	5.159	4.171	-1.179	0.000	5 THR	QG2 <==>	14 PRO	QG
2.73	2.462	2.445	-0.285	0.000	6 LYS	QB <==>	6 LYS	QD
3.64	3.423	2.979	-0.661	0.000	6 LYS	QB <==>	6 LYS	QE
2.74	2.488	2.467	-0.273	0.000	7 LYS	QB <==>	7 LYS	QD
2.53	3.185	2.787	0.257	0.257	7 LYS	QB <==>	7 LYS	QE
5.32	4.560	4.543	-0.777	0.000	9 PRO	QB <==>	10 GLN	HA
2.63	2.460	2.450	-0.180	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.771	4.756	-0.584	0.000	14 PRO	QB <==>	15 THR	HA

=====

SUMMARY: 105 unique NOE restraints displayed.

Violations (>0.001Å): 41 (39.0%)

Average violation (for violated restraints): 2.518 Å

Maximum violation: 8.476 Å

Violation severity distribution:

0-0.5Å: 9 restraints
>0.5Å: 8 restraints
>1.0Å: 6 restraints
>2.0Å: 7 restraints
>3.0Å: 11 restraints

4. DPCA04 (FOURTH CLUSTER) :

Starting NOE processing using 'carma.distances' at 2025-11-13 01:55:16
Defined 105 unique restraint groups from 257 calculated distances.
Processing 120,401 frames from 'carma.distances' in memory...
Progress: 120401/120401 (100.0%) | Elapsed: 11.0s | Remaining: 0.0s
Finished processing 120,401 frames.

NOE Analysis Results:

EXP	<r ⁻³ >	<r ⁻⁶ >	Deviation	Violation	Atoms Involved				
5.14	5.939	5.903	0.763	0.763	10	GLN	HB2 <==>	12	SER H
5.14	6.166	6.116	0.976	0.976	10	GLN	HB3 <==>	12	SER H
5.41	2.685	2.606	-2.804	0.000	11	VAL	HB <==>	12	SER H
5.50	3.158	3.073	-2.427	0.000	10	GLN	H <==>	10	GLN QG
5.50	6.653	6.608	1.108	1.108	10	GLN	QG <==>	12	SER H
5.50	5.910	5.877	0.377	0.377	5	THR	H <==>	13	THR QG2
5.50	2.662	2.624	-2.876	0.000	13	THR	H <==>	13	THR QG2
4.92	2.729	2.645	-2.275	0.000	15	THR	H <==>	15	THR QG2
5.50	3.010	2.879	-2.621	0.000	7	LYS	H <==>	7	LYS QG
5.50	4.799	4.734	-0.766	0.000	6	LYS	QG <==>	7	LYS H
5.50	4.273	4.042	-1.458	0.000	7	LYS	H <==>	7	LYS QD
5.50	2.943	2.866	-2.634	0.000	6	LYS	H <==>	6	LYS QG
5.50	4.252	3.720	-1.780	0.000	6	LYS	H <==>	6	LYS QD
5.50	5.093	4.881	-0.619	0.000	6	LYS	H <==>	6	LYS QE
3.75	4.608	4.308	0.558	0.558	7	LYS	HA <==>	7	LYS QE
3.08	4.600	4.178	1.098	1.098	6	LYS	HA <==>	6	LYS QE
2.97	2.457	2.445	-0.525	0.000	6	LYS	HB2 <==>	6	LYS QG
2.88	2.439	2.426	-0.454	0.000	7	LYS	HB2 <==>	7	LYS QG
5.50	4.324	4.086	-1.414	0.000	3	ARG	H <==>	3	ARG HG2
5.50	4.182	3.832	-1.668	0.000	3	ARG	H <==>	3	ARG HG3
4.33	4.373	3.900	-0.430	0.000	3	ARG	HA <==>	3	ARG QD
3.86	2.515	2.468	-1.392	0.000	7	LYS	HB3 <==>	8	VAL H
4.08	2.457	2.407	-1.673	0.000	8	VAL	H <==>	8	VAL QG2
5.15	3.968	3.770	-1.380	0.000	10	GLN	HB3 <==>	11	VAL H
3.11	2.135	2.122	-0.988	0.000	9	PRO	HA <==>	10	GLN H
3.18	2.358	2.343	-0.837	0.000	14	PRO	HA <==>	15	THR H
3.36	3.321	3.295	-0.065	0.000	12	SER	HA <==>	13	THR H
5.50	5.076	5.053	-0.447	0.000	8	VAL	H <==>	9	PRO HD2
5.50	2.566	2.495	-3.005	0.000	5	THR	H <==>	5	THR QG2
2.92	3.472	3.418	0.498	0.498	15	THR	HA <==>	16	LEU H
2.92	2.392	2.383	-0.537	0.000	13	THR	HA <==>	13	THR HB
5.50	2.794	2.745	-2.755	0.000	2	VAL	HB <==>	3	ARG H
5.50	4.447	4.440	-1.060	0.000	4	TYR	H <==>	5	THR H
2.86	2.558	2.526	-0.334	0.000	8	VAL	H <==>	8	VAL HB
4.28	3.034	2.661	-1.619	0.000	1	LEU	HA <==>	1	LEU QD1
4.28	2.491	2.349	-1.931	0.000	1	LEU	HA <==>	1	LEU QD2
4.65	11.841	11.604	6.954	6.954	1	LEU	QD1 <==>	4	TYR QD
4.65	11.767	11.623	6.973	6.973	1	LEU	QD2 <==>	4	TYR QD
5.50	12.875	12.679	7.179	7.179	1	LEU	QD1 <==>	4	TYR QE
5.50	12.883	12.716	7.216	7.216	1	LEU	QD2 <==>	4	TYR QE

4.77	3.920	3.804	-0.966	0.000	3 ARG HB3 <==>	4 TYR QE
3.65	2.526	2.501	-1.149	0.000	3 ARG HA <==>	3 ARG HG3
3.65	2.947	2.911	-0.739	0.000	3 ARG HA <==>	3 ARG HG2
4.77	5.141	4.875	0.105	0.105	3 ARG HB2 <==>	4 TYR QE
5.29	5.274	5.221	-0.069	0.000	3 ARG HA <==>	4 TYR QD
3.17	2.594	2.527	-0.643	0.000	4 TYR HA <==>	4 TYR QD
5.11	8.976	8.523	3.413	3.413	4 TYR QD <==>	9 PRO HD3
5.50	10.763	10.161	4.661	4.661	4 TYR QE <==>	9 PRO HD3
2.84	3.705	3.659	0.819	0.819	13 THR HA <==>	13 THR QG2
2.40	2.378	2.371	-0.029	0.000	7 LYS QG <==>	7 LYS QD
2.88	2.454	2.442	-0.438	0.000	7 LYS HB3 <==>	7 LYS QG
3.82	2.390	2.369	-1.451	0.000	7 LYS QG <==>	7 LYS QE
3.79	3.106	2.772	-1.018	0.000	7 LYS HA <==>	7 LYS QD
3.63	2.558	2.507	-1.123	0.000	7 LYS HA <==>	7 LYS QG
2.97	2.463	2.451	-0.519	0.000	6 LYS HB3 <==>	6 LYS QG
2.40	2.372	2.366	-0.034	0.000	6 LYS QG <==>	6 LYS QD
3.63	3.503	3.068	-0.562	0.000	6 LYS HA <==>	6 LYS QD
3.45	2.593	2.543	-0.907	0.000	6 LYS HA <==>	6 LYS QG
5.50	4.680	4.651	-0.849	0.000	7 LYS HA <==>	8 VAL HB
3.86	2.769	2.684	-1.176	0.000	10 GLN HA <==>	10 GLN QG
3.68	2.636	2.575	-1.105	0.000	11 VAL H <==>	11 VAL HB
4.56	4.136	3.999	-0.561	0.000	12 SER QB <==>	13 THR H
5.50	4.384	4.251	-1.249	0.000	7 LYS QD <==>	8 VAL H
3.86	3.757	3.720	-0.140	0.000	7 LYS HB2 <==>	8 VAL H
2.85	2.352	2.333	-0.517	0.000	7 LYS HA <==>	8 VAL H
5.50	4.312	4.273	-1.227	0.000	14 PRO HA <==>	16 LEU H
2.40	2.293	2.254	-0.146	0.000	8 VAL HA <==>	9 PRO HD3
2.41	2.367	2.324	-0.086	0.000	8 VAL HA <==>	9 PRO HD2
3.46	9.031	8.951	5.491	5.491	5 THR HA <==>	14 PRO HD2
3.46	10.217	10.138	6.678	6.678	5 THR HA <==>	14 PRO HD3
4.23	10.066	9.966	5.736	5.736	5 THR QG2 <==>	14 PRO HD2
4.11	4.028	3.604	-0.506	0.000	8 VAL HB <==>	9 PRO HD2
5.01	9.676	9.014	4.004	4.004	1 LEU QD1 <==>	9 PRO HA
5.01	9.699	9.200	4.190	4.190	1 LEU QD2 <==>	9 PRO HA
5.50	9.250	7.974	2.474	2.474	1 LEU QD2 <==>	9 PRO HD3
5.50	9.720	8.411	2.911	2.911	1 LEU QD1 <==>	9 PRO HD3
4.47	7.956	6.652	2.182	2.182	1 LEU QD1 <==>	9 PRO HB3
4.47	8.071	7.205	2.735	2.735	1 LEU QD2 <==>	9 PRO HB3
4.47	8.932	7.703	3.233	3.233	1 LEU QD1 <==>	9 PRO HB2
4.47	9.123	8.239	3.769	3.769	1 LEU QD2 <==>	9 PRO HB2
4.08	4.238	4.027	-0.053	0.000	8 VAL H <==>	8 VAL QG1
3.23	2.554	2.519	-0.711	0.000	8 VAL HA <==>	8 VAL QG2
3.23	2.557	2.533	-0.697	0.000	8 VAL HA <==>	8 VAL QG1
5.15	3.621	3.484	-1.666	0.000	10 GLN HB2 <==>	11 VAL H
5.50	10.861	10.662	5.162	5.162	4 TYR QD <==>	9 PRO HG2
5.14	12.406	11.442	6.302	6.302	4 TYR QE <==>	9 PRO HG2
5.20	4.721	4.598	-0.602	0.000	4 TYR QD <==>	13 THR QG2
4.23	11.462	11.373	7.143	7.143	5 THR QG2 <==>	14 PRO HD3
3.05	7.072	6.946	3.896	3.896	5 THR QG2 <==>	13 THR QG2
5.34	7.950	7.920	2.580	2.580	1 LEU HA <==>	3 ARG QB
2.24	2.402	2.396	0.156	0.156	3 ARG QB <==>	3 ARG QG
3.03	2.508	2.488	-0.542	0.000	3 ARG QB <==>	3 ARG QD
4.97	3.347	3.211	-1.759	0.000	3 ARG QB <==>	4 TYR QD
5.34	3.495	3.290	-2.050	0.000	3 ARG QG <==>	4 TYR QD
5.34	9.151	9.044	3.704	3.704	4 TYR QD <==>	9 PRO QB
4.33	8.215	8.166	3.836	3.836	5 THR H <==>	14 PRO QD
3.76	11.017	10.956	7.196	7.196	5 THR HB <==>	14 PRO QD
5.35	11.608	11.458	6.108	6.108	5 THR QG2 <==>	14 PRO QG
2.73	2.472	2.452	-0.278	0.000	6 LYS QB <==>	6 LYS QD
3.64	3.116	2.738	-0.902	0.000	6 LYS QB <==>	6 LYS QE
2.74	2.501	2.483	-0.257	0.000	7 LYS QB <==>	7 LYS QD
2.53	3.920	3.506	0.976	0.976	7 LYS QB <==>	7 LYS QE
5.32	5.069	5.052	-0.268	0.000	9 PRO QB <==>	10 GLN HA

2.63	2.436	2.424	-0.206	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.767	4.755	-0.585	0.000	14 PRO	QB <==>	15 THR	HA

```
=====
SUMMARY: 105 unique NOE restraints displayed.
Violations (>0.001Å): 37 (35.2%)
Average violation (for violated restraints): 3.599 Å
Maximum violation: 7.216 Å
Violation severity distribution:
  0-0.5Å: 4 restraints
  >0.5Å: 5 restraints
  >1.0Å: 2 restraints
  >2.0Å: 5 restraints
  >3.0Å: 21 restraints
```

5. DPCA05 (FIFTH CLUSTER):

Starting NOE processing using 'carma.distances' at 2025-11-13 01:56:01
 Defined 105 unique restraint groups from 257 calculated distances.
 Processing 133,577 frames from 'carma.distances' in memory...
 Progress: 133577/133577 (100.0%) | Elapsed: 12.0s | Remaining: 0.0s
 Finished processing 133,577 frames.

NOE Analysis Results:

EXP	<r ⁻³ >	<r ⁻⁶ >	Deviation	Violation	Atoms Involved			
5.14	6.219	6.178	1.038	1.038	10 GLN	HB2 <==>	12 SER	H
5.14	5.689	5.656	0.516	0.516	10 GLN	HB3 <==>	12 SER	H
5.41	2.930	2.780	-2.630	0.000	11 VAL	HB <==>	12 SER	H
5.50	2.682	2.623	-2.877	0.000	10 GLN	H <==>	10 GLN	QG
5.50	7.674	7.446	1.946	1.946	10 GLN	QG <==>	12 SER	H
5.50	8.485	8.317	2.817	2.817	5 THR	H <==>	13 THR	QG2
5.50	2.914	2.804	-2.696	0.000	13 THR	H <==>	13 THR	QG2
4.92	2.489	2.427	-2.493	0.000	15 THR	H <==>	15 THR	QG2
5.50	2.993	2.921	-2.579	0.000	7 LYS	H <==>	7 LYS	QG
5.50	4.814	4.781	-0.719	0.000	6 LYS	QG <==>	7 LYS	H
5.50	4.308	3.796	-1.704	0.000	7 LYS	H <==>	7 LYS	QD
5.50	2.973	2.881	-2.619	0.000	6 LYS	H <==>	6 LYS	QG
5.50	4.447	3.927	-1.573	0.000	6 LYS	H <==>	6 LYS	QD
5.50	5.216	5.027	-0.473	0.000	6 LYS	H <==>	6 LYS	QE
3.75	4.612	4.213	0.463	0.463	7 LYS	HA <==>	7 LYS	QE
3.08	4.821	4.512	1.432	1.432	6 LYS	HA <==>	6 LYS	QE
2.97	2.467	2.456	-0.514	0.000	6 LYS	HB2 <==>	6 LYS	QG
2.88	2.456	2.445	-0.435	0.000	7 LYS	HB2 <==>	7 LYS	QG
5.50	3.645	3.443	-2.057	0.000	3 ARG	H <==>	3 ARG	HG2
5.50	2.755	2.688	-2.812	0.000	3 ARG	H <==>	3 ARG	HG3
4.33	4.041	3.618	-0.712	0.000	3 ARG	HA <==>	3 ARG	QD
3.86	4.394	4.346	0.486	0.486	7 LYS	HB3 <==>	8 VAL	H
4.08	2.522	2.473	-1.607	0.000	8 VAL	H <==>	8 VAL	QG2
5.15	2.567	2.505	-2.645	0.000	10 GLN	HB3 <==>	11 VAL	H
3.11	2.240	2.227	-0.883	0.000	9 PRO	HA <==>	10 GLN	H
3.18	2.334	2.317	-0.863	0.000	14 PRO	HA <==>	15 THR	H
3.36	2.884	2.827	-0.533	0.000	12 SER	HA <==>	13 THR	H
5.50	5.057	5.045	-0.455	0.000	8 VAL	H <==>	9 PRO	HD2
5.50	3.904	3.742	-1.758	0.000	5 THR	H <==>	5 THR	QG2
2.92	3.301	3.155	0.235	0.235	15 THR	HA <==>	16 LEU	H
2.92	2.486	2.472	-0.448	0.000	13 THR	HA <==>	13 THR	HB

5.50	2.732	2.578	-2.922	0.000	2 VAL	HB <==>	3 ARG	H
5.50	4.382	4.372	-1.128	0.000	4 TYR	H <==>	5 THR	H
2.86	3.411	3.206	0.346	0.346	8 VAL	H <==>	8 VAL	HB
4.28	2.980	2.630	-1.650	0.000	1 LEU	HA <==>	1 LEU	QD1
4.28	2.501	2.351	-1.929	0.000	1 LEU	HA <==>	1 LEU	QD2
4.65	7.402	6.335	1.685	1.685	1 LEU	QD1 <==>	4 TYR	QD
4.65	7.604	6.777	2.127	2.127	1 LEU	QD2 <==>	4 TYR	QD
5.50	6.413	4.862	-0.638	0.000	1 LEU	QD1 <==>	4 TYR	QE
5.50	6.709	5.510	0.010	0.010	1 LEU	QD2 <==>	4 TYR	QE
4.77	7.603	7.447	2.677	2.677	3 ARG	HB3 <==>	4 TYR	QE
3.65	2.969	2.930	-0.720	0.000	3 ARG	HA <==>	3 ARG	HG3
3.65	2.584	2.549	-1.101	0.000	3 ARG	HA <==>	3 ARG	HG2
4.77	7.703	7.550	2.780	2.780	3 ARG	HB2 <==>	4 TYR	QE
5.29	3.666	3.457	-1.833	0.000	3 ARG	HA <==>	4 TYR	QD
3.17	3.128	3.066	-0.104	0.000	4 TYR	HA <==>	4 TYR	QD
5.11	3.476	3.323	-1.787	0.000	4 TYR	QD <==>	9 PRO	HD3
5.50	2.762	2.660	-2.840	0.000	4 TYR	QE <==>	9 PRO	HD3
2.84	2.496	2.474	-0.366	0.000	13 THR	HA <==>	13 THR	QG2
2.40	2.377	2.371	-0.029	0.000	7 LYS	QG <==>	7 LYS	QD
2.88	2.465	2.454	-0.426	0.000	7 LYS	HB3 <==>	7 LYS	QG
3.82	2.434	2.408	-1.412	0.000	7 LYS	QG <==>	7 LYS	QE
3.79	3.520	3.086	-0.704	0.000	7 LYS	HA <==>	7 LYS	QD
3.63	2.578	2.531	-1.099	0.000	7 LYS	HA <==>	7 LYS	QG
2.97	2.464	2.453	-0.517	0.000	6 LYS	HB3 <==>	6 LYS	QG
2.40	2.375	2.369	-0.031	0.000	6 LYS	QG <==>	6 LYS	QD
3.63	3.745	3.264	-0.366	0.000	6 LYS	HA <==>	6 LYS	QD
3.45	2.672	2.594	-0.856	0.000	6 LYS	HA <==>	6 LYS	QG
5.50	6.112	6.072	0.572	0.572	7 LYS	HA <==>	8 VAL	HB
3.86	2.520	2.495	-1.365	0.000	10 GLN	HA <==>	10 GLN	QG
3.68	3.180	3.008	-0.672	0.000	11 VAL	H <==>	11 VAL	HB
4.56	3.942	3.736	-0.824	0.000	12 SER	QB <==>	13 THR	H
5.50	6.125	5.998	0.498	0.498	7 LYS	QD <==>	8 VAL	H
3.86	4.159	4.123	0.263	0.263	7 LYS	HB2 <==>	8 VAL	H
2.85	3.037	3.018	0.168	0.168	7 LYS	HA <==>	8 VAL	H
5.50	4.379	4.337	-1.163	0.000	14 PRO	HA <==>	16 LEU	H
2.40	2.315	2.283	-0.117	0.000	8 VAL	HA <==>	9 PRO	HD3
2.41	2.525	2.475	0.065	0.065	8 VAL	HA <==>	9 PRO	HD2
3.46	12.156	11.949	8.489	8.489	5 THR	HA <==>	14 PRO	HD2
3.46	13.446	13.301	9.841	9.841	5 THR	HA <==>	14 PRO	HD3
4.23	10.052	9.645	5.415	5.415	5 THR	QG2 <==>	14 PRO	HD2
4.11	2.451	2.362	-1.748	0.000	8 VAL	HB <==>	9 PRO	HD2
5.01	10.176	9.441	4.431	4.431	1 LEU	QD1 <==>	9 PRO	HA
5.01	10.277	9.847	4.837	4.837	1 LEU	QD2 <==>	9 PRO	HA
5.50	11.843	10.237	4.737	4.737	1 LEU	QD2 <==>	9 PRO	HD3
5.50	11.946	9.848	4.348	4.348	1 LEU	QD1 <==>	9 PRO	HD3
4.47	9.378	7.560	3.090	3.090	1 LEU	QD1 <==>	9 PRO	HB3
4.47	9.471	8.406	3.936	3.936	1 LEU	QD2 <==>	9 PRO	HB3
4.47	10.286	7.926	3.456	3.456	1 LEU	QD1 <==>	9 PRO	HB2
4.47	10.343	9.015	4.545	4.545	1 LEU	QD2 <==>	9 PRO	HB2
4.08	3.424	3.323	-0.757	0.000	8 VAL	H <==>	8 VAL	QG1
3.23	3.312	3.120	-0.110	0.000	8 VAL	HA <==>	8 VAL	QG2
3.23	2.444	2.422	-0.808	0.000	8 VAL	HA <==>	8 VAL	QG1
5.15	3.756	3.695	-1.455	0.000	10 GLN	HB2 <==>	11 VAL	H
5.50	5.457	5.318	-0.182	0.000	4 TYR	QD <==>	9 PRO	HG2
5.14	4.991	4.860	-0.280	0.000	4 TYR	QE <==>	9 PRO	HG2
5.20	10.026	9.633	4.433	4.433	4 TYR	QD <==>	13 THR	QG2
4.23	11.305	10.990	6.760	6.760	5 THR	QG2 <==>	14 PRO	HD3
3.05	7.376	6.818	3.768	3.768	5 THR	QG2 <==>	13 THR	QG2
5.34	8.382	8.311	2.971	2.971	1 LEU	HA <==>	3 ARG	QB
2.24	2.390	2.384	0.144	0.144	3 ARG	QB <==>	3 ARG	QG
3.03	2.409	2.397	-0.633	0.000	3 ARG	QB <==>	3 ARG	QD
4.97	5.559	5.437	0.467	0.467	3 ARG	QB <==>	4 TYR	QD
5.34	6.182	6.005	0.665	0.665	3 ARG	QG <==>	4 TYR	QD

5.34	3.499	3.248	-2.092	0.000	4 TYR	QD <==>	9 PRO	QB
4.33	11.369	11.265	6.935	6.935	5 THR	H <==>	14 PRO	QD
3.76	11.409	11.186	7.426	7.426	5 THR	HB <==>	14 PRO	QD
5.35	11.748	11.138	5.788	5.788	5 THR	QG2 <==>	14 PRO	QG
2.73	2.459	2.442	-0.288	0.000	6 LYS	QB <==>	6 LYS	QD
3.64	3.152	2.761	-0.879	0.000	6 LYS	QB <==>	6 LYS	QE
2.74	2.468	2.449	-0.291	0.000	7 LYS	QB <==>	7 LYS	QD
2.53	3.150	2.760	0.230	0.230	7 LYS	QB <==>	7 LYS	QE
5.32	4.699	4.678	-0.642	0.000	9 PRO	QB <==>	10 GLN	HA
2.63	2.488	2.481	-0.149	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.948	4.930	-0.410	0.000	14 PRO	QB <==>	15 THR	HA

=====

SUMMARY: 105 unique NOE restraints displayed.

Violations (>0.001Å): 41 (39.0%)

Average violation (for violated restraints): 2.850 Å

Maximum violation: 9.841 Å

Violation severity distribution:

- 0-0.5Å: 12 restraints
- >0.5Å: 3 restraints
- >1.0Å: 4 restraints
- >2.0Å: 5 restraints
- >3.0Å: 17 restraints

6. DPCA06 (SIXTH CLUSTER) :

Starting NOE processing using 'carma.distances' at 2025-11-13 01:56:59

Defined 105 unique restraint groups from 257 calculated distances.

Processing 123,575 frames from 'carma.distances' in memory...

Progress: 123575/123575 (100.0%) | Elapsed: 11.1s | Remaining: 0.0s

Finished processing 123,575 frames.

NOE Analysis Results:

EXP	<r^-3>	<r^-6>	Deviation	Violation	Atoms Involved			
5.14	5.358	5.317	0.177	0.177	10 GLN	HB2 <==>	12 SER	H
5.14	5.554	5.506	0.366	0.366	10 GLN	HB3 <==>	12 SER	H
5.41	4.022	3.910	-1.500	0.000	11 VAL	HB <==>	12 SER	H
5.50	3.157	3.052	-2.448	0.000	10 GLN	H <==>	10 GLN	QG
5.50	6.135	6.093	0.593	0.593	10 GLN	QG <==>	12 SER	H
5.50	9.695	9.015	3.515	3.515	5 THR	H <==>	13 THR	QG2
5.50	3.368	3.049	-2.451	0.000	13 THR	H <==>	13 THR	QG2
4.92	3.292	3.149	-1.771	0.000	15 THR	H <==>	15 THR	QG2
5.50	3.094	2.954	-2.546	0.000	7 LYS	H <==>	7 LYS	QG
5.50	3.598	3.280	-2.220	0.000	6 LYS	QG <==>	7 LYS	H
5.50	4.616	4.263	-1.237	0.000	7 LYS	H <==>	7 LYS	QD
5.50	3.025	2.891	-2.609	0.000	6 LYS	H <==>	6 LYS	QG
5.50	4.234	3.636	-1.864	0.000	6 LYS	H <==>	6 LYS	QD
5.50	4.668	4.342	-1.158	0.000	6 LYS	H <==>	6 LYS	QE
3.75	4.970	4.884	1.134	1.134	7 LYS	HA <==>	7 LYS	QE
3.08	4.892	4.663	1.583	1.583	6 LYS	HA <==>	6 LYS	QE
2.97	2.474	2.460	-0.510	0.000	6 LYS	HB2 <==>	6 LYS	QG
2.88	2.477	2.466	-0.414	0.000	7 LYS	HB2 <==>	7 LYS	QG
5.50	3.211	2.911	-2.589	0.000	3 ARG	H <==>	3 ARG	HG2
5.50	3.279	3.047	-2.453	0.000	3 ARG	H <==>	3 ARG	HG3
4.33	3.679	3.197	-1.133	0.000	3 ARG	HA <==>	3 ARG	QD
3.86	2.986	2.870	-0.990	0.000	7 LYS	HB3 <==>	8 VAL	H
4.08	2.306	2.264	-1.816	0.000	8 VAL	H <==>	8 VAL	QG2

5.15	3.553	3.362	-1.788	0.000	10 GLN	HB3 <==>	11 VAL	H
3.11	2.239	2.223	-0.887	0.000	9 PRO	HA <==>	10 GLN	H
3.18	2.233	2.219	-0.961	0.000	14 PRO	HA <==>	15 THR	H
3.36	3.088	3.064	-0.296	0.000	12 SER	HA <==>	13 THR	H
5.50	5.170	5.153	-0.347	0.000	8 VAL	H <==>	9 PRO	HD2
5.50	3.098	2.965	-2.535	0.000	5 THR	H <==>	5 THR	QG2
2.92	2.280	2.261	-0.659	0.000	15 THR	HA <==>	16 LEU	H
2.92	2.809	2.767	-0.153	0.000	13 THR	HA <==>	13 THR	HB
5.50	2.803	2.645	-2.855	0.000	2 VAL	HB <==>	3 ARG	H
5.50	4.459	4.348	-1.152	0.000	4 TYR	H <==>	5 THR	H
2.86	2.641	2.581	-0.279	0.000	8 VAL	H <==>	8 VAL	HB
4.28	2.840	2.557	-1.723	0.000	1 LEU	HA <==>	1 LEU	QD1
4.28	2.652	2.441	-1.839	0.000	1 LEU	HA <==>	1 LEU	QD2
4.65	6.907	5.247	0.597	0.597	1 LEU	QD1 <==>	4 TYR	QD
4.65	7.418	5.830	1.180	1.180	1 LEU	QD2 <==>	4 TYR	QD
5.50	5.948	4.403	-1.097	0.000	1 LEU	QD1 <==>	4 TYR	QE
5.50	6.335	4.689	-0.811	0.000	1 LEU	QD2 <==>	4 TYR	QE
4.77	7.182	6.918	2.148	2.148	3 ARG	HB3 <==>	4 TYR	QE
3.65	3.074	2.936	-0.714	0.000	3 ARG	HA <==>	3 ARG	HG3
3.65	2.953	2.821	-0.829	0.000	3 ARG	HA <==>	3 ARG	HG2
4.77	7.217	6.874	2.104	2.104	3 ARG	HB2 <==>	4 TYR	QE
5.29	3.947	3.732	-1.558	0.000	3 ARG	HA <==>	4 TYR	QD
3.17	2.712	2.600	-0.570	0.000	4 TYR	HA <==>	4 TYR	QD
5.11	11.170	10.174	5.064	5.064	4 TYR	QD <==>	9 PRO	HD3
5.50	11.768	9.531	4.031	4.031	4 TYR	QE <==>	9 PRO	HD3
2.84	2.478	2.454	-0.386	0.000	13 THR	HA <==>	13 THR	QG2
2.40	2.370	2.363	-0.037	0.000	7 LYS	QG <==>	7 LYS	QD
2.88	2.460	2.449	-0.431	0.000	7 LYS	HB3 <==>	7 LYS	QG
3.82	2.424	2.398	-1.422	0.000	7 LYS	QG <==>	7 LYS	QE
3.79	4.188	3.836	0.046	0.046	7 LYS	HA <==>	7 LYS	QD
3.63	2.650	2.585	-1.045	0.000	7 LYS	HA <==>	7 LYS	QG
2.97	2.452	2.440	-0.530	0.000	6 LYS	HB3 <==>	6 LYS	QG
2.40	2.367	2.360	-0.040	0.000	6 LYS	QG <==>	6 LYS	QD
3.63	4.024	3.577	-0.053	0.000	6 LYS	HA <==>	6 LYS	QD
3.45	2.675	2.608	-0.842	0.000	6 LYS	HA <==>	6 LYS	QG
5.50	4.821	4.773	-0.727	0.000	7 LYS	HA <==>	8 VAL	HB
3.86	2.633	2.574	-1.286	0.000	10 GLN	HA <==>	10 GLN	QG
3.68	3.513	3.367	-0.313	0.000	11 VAL	H <==>	11 VAL	HB
4.56	3.942	3.794	-0.766	0.000	12 SER	QB <==>	13 THR	H
5.50	5.013	4.916	-0.584	0.000	7 LYS	QD <==>	8 VAL	H
3.86	3.502	3.258	-0.602	0.000	7 LYS	HB2 <==>	8 VAL	H
2.85	2.223	2.209	-0.641	0.000	7 LYS	HA <==>	8 VAL	H
5.50	6.373	6.363	0.863	0.863	14 PRO	HA <==>	16 LEU	H
2.40	2.239	2.205	-0.195	0.000	8 VAL	HA <==>	9 PRO	HD3
2.41	2.494	2.440	0.030	0.030	8 VAL	HA <==>	9 PRO	HD2
3.46	10.014	9.326	5.866	5.866	5 THR	HA <==>	14 PRO	HD2
3.46	9.652	9.033	5.573	5.573	5 THR	HA <==>	14 PRO	HD3
4.23	11.407	10.996	6.766	6.766	5 THR	QG2 <==>	14 PRO	HD2
4.11	3.472	3.022	-1.088	0.000	8 VAL	HB <==>	9 PRO	HD2
5.01	20.655	18.757	13.747	13.747	1 LEU	QD1 <==>	9 PRO	HA
5.01	20.888	19.350	14.340	14.340	1 LEU	QD2 <==>	9 PRO	HA
5.50	20.094	18.183	12.683	12.683	1 LEU	QD2 <==>	9 PRO	HD3
5.50	19.880	17.651	12.151	12.151	1 LEU	QD1 <==>	9 PRO	HD3
4.47	22.400	20.243	15.773	15.773	1 LEU	QD1 <==>	9 PRO	HB3
4.47	22.623	20.945	16.475	16.475	1 LEU	QD2 <==>	9 PRO	HB3
4.47	22.785	20.726	16.256	16.256	1 LEU	QD1 <==>	9 PRO	HB2
4.47	22.990	21.408	16.938	16.938	1 LEU	QD2 <==>	9 PRO	HB2
4.08	3.625	3.192	-0.888	0.000	8 VAL	H <==>	8 VAL	QG1
3.23	2.655	2.593	-0.637	0.000	8 VAL	HA <==>	8 VAL	QG2
3.23	2.550	2.523	-0.707	0.000	8 VAL	HA <==>	8 VAL	QG1
5.15	3.251	3.116	-2.034	0.000	10 GLN	HB2 <==>	11 VAL	H
5.50	13.695	13.043	7.543	7.543	4 TYR	QD <==>	9 PRO	HG2
5.14	14.155	11.593	6.453	6.453	4 TYR	QE <==>	9 PRO	HG2

5.20	7.354	6.500	1.300	1.300	4 TYR	QD <==>	13 THR	QG2
4.23	11.305	10.804	6.574	6.574	5 THR	QG2 <==>	14 PRO	HD3
3.05	12.291	10.610	7.560	7.560	5 THR	QG2 <==>	13 THR	QG2
5.34	7.216	6.962	1.622	1.622	1 LEU	HA <==>	3 ARG	QB
2.24	2.372	2.364	0.124	0.124	3 ARG	QB <==>	3 ARG	QG
3.03	2.481	2.455	-0.575	0.000	3 ARG	QB <==>	3 ARG	QD
4.97	5.434	5.136	0.166	0.166	3 ARG	QB <==>	4 TYR	QD
5.34	5.933	5.514	0.174	0.174	3 ARG	QG <==>	4 TYR	QD
5.34	13.889	13.501	8.161	8.161	4 TYR	QD <==>	9 PRO	QB
4.33	8.561	8.250	3.920	3.920	5 THR	H <==>	14 PRO	QD
3.76	11.153	10.732	6.972	6.972	5 THR	HB <==>	14 PRO	QD
5.35	9.637	8.816	3.466	3.466	5 THR	QG2 <==>	14 PRO	QG
2.73	2.475	2.454	-0.276	0.000	6 LYS	QB <==>	6 LYS	QD
3.64	2.726	2.461	-1.179	0.000	6 LYS	QB <==>	6 LYS	QE
2.74	2.460	2.444	-0.296	0.000	7 LYS	QB <==>	7 LYS	QD
2.53	3.096	2.728	0.198	0.198	7 LYS	QB <==>	7 LYS	QE
5.32	4.756	4.731	-0.589	0.000	9 PRO	QB <==>	10 GLN	HA
2.63	2.440	2.428	-0.202	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.444	4.415	-0.925	0.000	14 PRO	QB <==>	15 THR	HA

=====

SUMMARY: 105 unique NOE restraints displayed.
 Violations (>0.001Å): 40 (38.1%)
 Average violation (for violated restraints): 5.356 Å
 Maximum violation: 16.938 Å
 Violation severity distribution:
 0-0.5Å: 8 restraints
 >0.5Å: 3 restraints
 >1.0Å: 5 restraints
 >2.0Å: 2 restraints
 >3.0Å: 22 restraints

7. DPCA07 (SEVENTH CLUSTER):

Starting NOE processing using 'carma.distances' at 2025-11-13 01:57:47
 Defined 105 unique restraint groups from 257 calculated distances.
 Processing 159,039 frames from 'carma.distances' in memory...
 Progress: 159039/159039 (100.0%) | Elapsed: 14.4s | Remaining: 0.0s
 Finished processing 159,039 frames.

NOE Analysis Results:

EXP	<r^-3>	<r^-6>	Deviation	Violation	Atoms Involved			
5.14	6.529	5.688	0.548	0.548	10 GLN	HB2 <==>	12 SER	H
5.14	6.187	4.980	-0.160	0.000	10 GLN	HB3 <==>	12 SER	H
5.41	2.956	2.803	-2.607	0.000	11 VAL	HB <==>	12 SER	H
5.50	2.855	2.691	-2.809	0.000	10 GLN	H <==>	10 GLN	QG
5.50	6.770	4.925	-0.575	0.000	10 GLN	QG <==>	12 SER	H
5.50	10.169	9.615	4.115	4.115	5 THR	H <==>	13 THR	QG2
5.50	3.305	3.089	-2.411	0.000	13 THR	H <==>	13 THR	QG2
4.92	3.080	3.001	-1.919	0.000	15 THR	H <==>	15 THR	QG2
5.50	2.730	2.558	-2.942	0.000	7 LYS	H <==>	7 LYS	QG
5.50	3.605	3.300	-2.200	0.000	6 LYS	QG <==>	7 LYS	H
5.50	4.214	3.981	-1.519	0.000	7 LYS	H <==>	7 LYS	QD

5.50	2.901	2.770	-2.730	0.000	6 LYS	H <==>	6 LYS	QG
5.50	4.544	4.298	-1.202	0.000	6 LYS	H <==>	6 LYS	QD
5.50	5.233	5.044	-0.456	0.000	6 LYS	H <==>	6 LYS	QE
3.75	4.764	4.402	0.652	0.652	7 LYS	HA <==>	7 LYS	QE
3.08	4.713	4.358	1.278	1.278	6 LYS	HA <==>	6 LYS	QE
2.97	2.458	2.447	-0.523	0.000	6 LYS	HB2 <==>	6 LYS	QG
2.88	2.461	2.449	-0.431	0.000	7 LYS	HB2 <==>	7 LYS	QG
5.50	3.683	3.368	-2.132	0.000	3 ARG	H <==>	3 ARG	HG2
5.50	3.123	3.024	-2.476	0.000	3 ARG	H <==>	3 ARG	HG3
4.33	3.838	3.325	-1.005	0.000	3 ARG	HA <==>	3 ARG	QD
3.86	3.148	2.969	-0.891	0.000	7 LYS	HB3 <==>	8 VAL	H
4.08	2.778	2.663	-1.417	0.000	8 VAL	H <==>	8 VAL	QG2
5.15	3.768	3.566	-1.584	0.000	10 GLN	HB3 <==>	11 VAL	H
3.11	2.430	2.355	-0.755	0.000	9 PRO	HA <==>	10 GLN	H
3.18	2.280	2.262	-0.918	0.000	14 PRO	HA <==>	15 THR	H
3.36	3.271	3.201	-0.159	0.000	12 SER	HA <==>	13 THR	H
5.50	3.617	3.424	-2.076	0.000	8 VAL	H <==>	9 PRO	HD2
5.50	2.684	2.595	-2.905	0.000	5 THR	H <==>	5 THR	QG2
2.92	3.333	3.298	0.378	0.378	15 THR	HA <==>	16 LEU	H
2.92	2.653	2.619	-0.301	0.000	13 THR	HA <==>	13 THR	HB
5.50	3.532	3.187	-2.313	0.000	2 VAL	HB <==>	3 ARG	H
5.50	3.903	3.351	-2.149	0.000	4 TYR	H <==>	5 THR	H
2.86	3.759	3.687	0.827	0.827	8 VAL	H <==>	8 VAL	HB
4.28	3.183	2.770	-1.510	0.000	1 LEU	HA <==>	1 LEU	QD1
4.28	2.490	2.363	-1.917	0.000	1 LEU	HA <==>	1 LEU	QD2
4.65	6.776	5.112	0.462	0.462	1 LEU	QD1 <==>	4 TYR	QD
4.65	7.073	5.151	0.501	0.501	1 LEU	QD2 <==>	4 TYR	QD
5.50	5.950	4.290	-1.210	0.000	1 LEU	QD1 <==>	4 TYR	QE
5.50	6.575	4.670	-0.830	0.000	1 LEU	QD2 <==>	4 TYR	QE
4.77	6.252	5.487	0.717	0.717	3 ARG	HB3 <==>	4 TYR	QE
3.65	2.995	2.934	-0.716	0.000	3 ARG	HA <==>	3 ARG	HG3
3.65	2.629	2.574	-1.076	0.000	3 ARG	HA <==>	3 ARG	HG2
4.77	6.604	5.666	0.896	0.896	3 ARG	HB2 <==>	4 TYR	QE
5.29	3.341	3.195	-2.095	0.000	3 ARG	HA <==>	4 TYR	QD
3.17	2.894	2.815	-0.355	0.000	4 TYR	HA <==>	4 TYR	QD
5.11	5.364	4.990	-0.120	0.000	4 TYR	QD <==>	9 PRO	HD3
5.50	6.397	5.988	0.488	0.488	4 TYR	QE <==>	9 PRO	HD3
2.84	2.505	2.471	-0.369	0.000	13 THR	HA <==>	13 THR	QG2
2.40	2.373	2.367	-0.033	0.000	7 LYS	QG <==>	7 LYS	QD
2.88	2.467	2.455	-0.425	0.000	7 LYS	HB3 <==>	7 LYS	QG
3.82	2.439	2.412	-1.408	0.000	7 LYS	QG <==>	7 LYS	QE
3.79	3.687	3.209	-0.581	0.000	7 LYS	HA <==>	7 LYS	QD
3.63	2.658	2.587	-1.043	0.000	7 LYS	HA <==>	7 LYS	QG
2.97	2.470	2.459	-0.511	0.000	6 LYS	HB3 <==>	6 LYS	QG
2.40	2.369	2.362	-0.038	0.000	6 LYS	QG <==>	6 LYS	QD
3.63	3.636	3.144	-0.486	0.000	6 LYS	HA <==>	6 LYS	QD
3.45	2.778	2.673	-0.777	0.000	6 LYS	HA <==>	6 LYS	QG
5.50	5.548	5.492	-0.008	0.000	7 LYS	HA <==>	8 VAL	HB
3.86	2.672	2.600	-1.260	0.000	10 GLN	HA <==>	10 GLN	QG
3.68	3.324	3.150	-0.530	0.000	11 VAL	H <==>	11 VAL	HB
4.56	3.581	3.409	-1.151	0.000	12 SER	QB <==>	13 THR	H
5.50	4.704	4.243	-1.257	0.000	7 LYS	QD <==>	8 VAL	H
3.86	3.546	3.305	-0.555	0.000	7 LYS	HB2 <==>	8 VAL	H
2.85	2.230	2.206	-0.644	0.000	7 LYS	HA <==>	8 VAL	H
5.50	4.079	4.051	-1.449	0.000	14 PRO	HA <==>	16 LEU	H
2.40	2.567	2.471	0.071	0.071	8 VAL	HA <==>	9 PRO	HD3
2.41	2.511	2.449	0.039	0.039	8 VAL	HA <==>	9 PRO	HD2
3.46	9.679	8.575	5.115	5.115	5 THR	HA <==>	14 PRO	HD2
3.46	9.418	8.299	4.839	4.839	5 THR	HA <==>	14 PRO	HD3
4.23	10.175	9.393	5.163	5.163	5 THR	QG2 <==>	14 PRO	HD2
4.11	3.170	2.760	-1.350	0.000	8 VAL	HB <==>	9 PRO	HD2
5.01	12.888	11.439	6.429	6.429	1 LEU	QD1 <==>	9 PRO	HA
5.01	13.322	12.100	7.090	7.090	1 LEU	QD2 <==>	9 PRO	HA

5.50	13.218	11.319	5.819	5.819	1 LEU	QD2 <==>	9 PRO	HD3
5.50	12.968	11.201	5.701	5.701	1 LEU	QD1 <==>	9 PRO	HD3
4.47	11.782	9.897	5.427	5.427	1 LEU	QD1 <==>	9 PRO	HB3
4.47	12.136	10.088	5.618	5.618	1 LEU	QD2 <==>	9 PRO	HB3
4.47	11.345	9.133	4.663	4.663	1 LEU	QD1 <==>	9 PRO	HB2
4.47	11.808	9.509	5.039	5.039	1 LEU	QD2 <==>	9 PRO	HB2
4.08	3.068	2.885	-1.195	0.000	8 VAL	H <==>	8 VAL	QG1
3.23	3.169	3.000	-0.230	0.000	8 VAL	HA <==>	8 VAL	QG2
3.23	2.402	2.382	-0.848	0.000	8 VAL	HA <==>	8 VAL	QG1
5.15	3.474	3.336	-1.814	0.000	10 GLN	HB2 <==>	11 VAL	H
5.50	4.229	3.748	-1.752	0.000	4 TYR	QD <==>	9 PRO	HG2
5.14	4.467	4.077	-1.063	0.000	4 TYR	QE <==>	9 PRO	HG2
5.20	7.538	5.795	0.595	0.595	4 TYR	QD <==>	13 THR	QG2
4.23	10.173	9.613	5.383	5.383	5 THR	QG2 <==>	14 PRO	HD3
3.05	11.113	9.790	6.740	6.740	5 THR	QG2 <==>	13 THR	QG2
5.34	5.942	5.766	0.426	0.426	1 LEU	HA <==>	3 ARG	QB
2.24	2.378	2.371	0.131	0.131	3 ARG	QB <==>	3 ARG	QG
3.03	2.483	2.461	-0.569	0.000	3 ARG	QB <==>	3 ARG	QD
4.97	4.718	4.395	-0.575	0.000	3 ARG	QB <==>	4 TYR	QD
5.34	5.118	4.736	-0.604	0.000	3 ARG	QG <==>	4 TYR	QD
5.34	3.681	3.285	-2.055	0.000	4 TYR	QD <==>	9 PRO	QB
4.33	8.875	8.688	4.358	4.358	5 THR	H <==>	14 PRO	QD
3.76	10.589	9.494	5.734	5.734	5 THR	HB <==>	14 PRO	QD
5.35	8.724	7.660	2.310	2.310	5 THR	QG2 <==>	14 PRO	QG
2.73	2.479	2.460	-0.270	0.000	6 LYS	QB <==>	6 LYS	QD
3.64	2.950	2.620	-1.020	0.000	6 LYS	QB <==>	6 LYS	QE
2.74	2.468	2.450	-0.290	0.000	7 LYS	QB <==>	7 LYS	QD
2.53	3.102	2.730	0.200	0.200	7 LYS	QB <==>	7 LYS	QE
5.32	4.605	4.575	-0.745	0.000	9 PRO	QB <==>	10 GLN	HA
2.63	2.457	2.445	-0.185	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.629	4.611	-0.729	0.000	14 PRO	QB <==>	15 THR	HA

=====

SUMMARY: 105 unique NOE restraints displayed.

Violations (>0.001Å): 33 (31.4%)

Average violation (for violated restraints): 2.962 Å

Maximum violation: 7.090 Å

Violation severity distribution:

- 0-0.5Å: 8 restraints
- >0.5Å: 7 restraints
- >1.0Å: 1 restraints
- >2.0Å: 1 restraints
- >3.0Å: 16 restraints

8. DPCA08 (EIGHTH CLUSTER):

Starting NOE processing using 'carma.distances' at 2025-11-13 01:58:40

Defined 105 unique restraint groups from 257 calculated distances.

Processing 73,161 frames from 'carma.distances' in memory...

Progress: 73161/73161 (100.0%) | Elapsed: 6.7s | Remaining: 0.0s

Finished processing 73,161 frames.

NOE Analysis Results:								
EXP	<r^-3>	<r^-6>	Deviation	Violation	Atoms Involved			
5.14	4.653	4.598	-0.542	0.000	10 GLN	HB2 <==>	12 SER	H
5.14	3.252	3.196	-1.944	0.000	10 GLN	HB3 <==>	12 SER	H
5.41	3.896	3.768	-1.642	0.000	11 VAL	HB <==>	12 SER	H
5.50	3.131	3.094	-2.406	0.000	10 GLN	H <==>	10 GLN	QG

5.50	5.490	5.225	-0.275	0.000	10 GLN	QG <==>	12 SER	H
5.50	5.788	5.729	0.229	0.229	5 THR	H <==>	13 THR	QG2
5.50	2.645	2.594	-2.906	0.000	13 THR	H <==>	13 THR	QG2
4.92	2.740	2.619	-2.301	0.000	15 THR	H <==>	15 THR	QG2
5.50	3.264	3.059	-2.441	0.000	7 LYS	H <==>	7 LYS	QG
5.50	4.008	3.751	-1.749	0.000	6 LYS	QG <==>	7 LYS	H
5.50	4.678	4.553	-0.947	0.000	7 LYS	H <==>	7 LYS	QD
5.50	2.906	2.615	-2.885	0.000	6 LYS	H <==>	6 LYS	QG
5.50	4.334	4.175	-1.325	0.000	6 LYS	H <==>	6 LYS	QD
5.50	5.089	4.834	-0.666	0.000	6 LYS	H <==>	6 LYS	QE
3.75	4.467	3.971	0.221	0.221	7 LYS	HA <==>	7 LYS	QE
3.08	4.436	3.899	0.819	0.819	6 LYS	HA <==>	6 LYS	QE
2.97	2.465	2.454	-0.516	0.000	6 LYS	HB2 <==>	6 LYS	QG
2.88	2.458	2.447	-0.433	0.000	7 LYS	HB2 <==>	7 LYS	QG
5.50	3.512	3.374	-2.126	0.000	3 ARG	H <==>	3 ARG	HG2
5.50	2.499	2.453	-3.047	0.000	3 ARG	H <==>	3 ARG	HG3
4.33	4.268	4.063	-0.267	0.000	3 ARG	HA <==>	3 ARG	QD
3.86	3.181	3.027	-0.833	0.000	7 LYS	HB3 <==>	8 VAL	H
4.08	2.590	2.543	-1.537	0.000	8 VAL	H <==>	8 VAL	QG2
5.15	2.613	2.574	-2.576	0.000	10 GLN	HB3 <==>	11 VAL	H
3.11	2.275	2.261	-0.849	0.000	9 PRO	HA <==>	10 GLN	H
3.18	2.355	2.340	-0.840	0.000	14 PRO	HA <==>	15 THR	H
3.36	3.230	3.207	-0.153	0.000	12 SER	HA <==>	13 THR	H
5.50	5.038	5.023	-0.477	0.000	8 VAL	H <==>	9 PRO	HD2
5.50	2.746	2.691	-2.809	0.000	5 THR	H <==>	5 THR	QG2
2.92	2.248	2.234	-0.686	0.000	15 THR	HA <==>	16 LEU	H
2.92	2.524	2.515	-0.405	0.000	13 THR	HA <==>	13 THR	HB
5.50	2.393	2.358	-3.142	0.000	2 VAL	HB <==>	3 ARG	H
5.50	2.887	2.867	-2.633	0.000	4 TYR	H <==>	5 THR	H
2.86	3.608	3.470	0.610	0.610	8 VAL	H <==>	8 VAL	HB
4.28	3.456	2.974	-1.306	0.000	1 LEU	HA <==>	1 LEU	QD1
4.28	2.352	2.269	-2.011	0.000	1 LEU	HA <==>	1 LEU	QD2
4.65	11.674	11.388	6.738	6.738	1 LEU	QD1 <==>	4 TYR	QD
4.65	11.640	11.402	6.752	6.752	1 LEU	QD2 <==>	4 TYR	QD
5.50	12.759	12.557	7.057	7.057	1 LEU	QD1 <==>	4 TYR	QE
5.50	12.904	12.678	7.178	7.178	1 LEU	QD2 <==>	4 TYR	QE
4.77	3.574	3.504	-1.266	0.000	3 ARG	HB3 <==>	4 TYR	QE
3.65	3.002	2.963	-0.687	0.000	3 ARG	HA <==>	3 ARG	HG3
3.65	2.510	2.485	-1.165	0.000	3 ARG	HA <==>	3 ARG	HG2
4.77	3.602	3.487	-1.283	0.000	3 ARG	HB2 <==>	4 TYR	QE
5.29	4.884	4.847	-0.443	0.000	3 ARG	HA <==>	4 TYR	QD
3.17	2.500	2.465	-0.705	0.000	4 TYR	HA <==>	4 TYR	QD
5.11	12.124	12.028	6.918	6.918	4 TYR	QD <==>	9 PRO	HD3
5.50	14.511	14.333	8.833	8.833	4 TYR	QE <==>	9 PRO	HD3
2.84	2.441	2.424	-0.416	0.000	13 THR	HA <==>	13 THR	QG2
2.40	2.374	2.367	-0.033	0.000	7 LYS	QG <==>	7 LYS	QD
2.88	2.462	2.452	-0.428	0.000	7 LYS	HB3 <==>	7 LYS	QG
3.82	2.437	2.412	-1.408	0.000	7 LYS	QG <==>	7 LYS	QE
3.79	3.387	2.961	-0.829	0.000	7 LYS	HA <==>	7 LYS	QD
3.63	2.597	2.546	-1.084	0.000	7 LYS	HA <==>	7 LYS	QG
2.97	2.450	2.439	-0.531	0.000	6 LYS	HB3 <==>	6 LYS	QG
2.40	2.376	2.369	-0.031	0.000	6 LYS	QG <==>	6 LYS	QD
3.63	3.349	2.933	-0.697	0.000	6 LYS	HA <==>	6 LYS	QD
3.45	2.622	2.561	-0.889	0.000	6 LYS	HA <==>	6 LYS	QG
5.50	6.293	6.255	0.755	0.755	7 LYS	HA <==>	8 VAL	HB
3.86	3.012	2.944	-0.916	0.000	10 GLN	HA <==>	10 GLN	QG
3.68	3.650	3.545	-0.135	0.000	11 VAL	H <==>	11 VAL	HB
4.56	3.949	3.872	-0.688	0.000	12 SER	QB <==>	13 THR	H
5.50	5.276	5.214	-0.286	0.000	7 LYS	QD <==>	8 VAL	H
3.86	3.211	3.065	-0.795	0.000	7 LYS	HB2 <==>	8 VAL	H
2.85	3.515	3.508	0.658	0.658	7 LYS	HA <==>	8 VAL	H
5.50	6.360	6.353	0.853	0.853	14 PRO	HA <==>	16 LEU	H
2.40	2.301	2.265	-0.135	0.000	8 VAL	HA <==>	9 PRO	HD3

2.41	2.516	2.466	0.056	0.056	8 VAL	HA <==>	9 PRO	HD2
3.46	9.206	9.152	5.692	5.692	5 THR	HA <==>	14 PRO	HD2
3.46	10.333	10.274	6.814	6.814	5 THR	HA <==>	14 PRO	HD3
4.23	7.484	7.257	3.027	3.027	5 THR	QG2 <==>	14 PRO	HD2
4.11	2.493	2.412	-1.698	0.000	8 VAL	HB <==>	9 PRO	HD2
5.01	10.109	9.346	4.336	4.336	1 LEU	QD1 <==>	9 PRO	HA
5.01	9.221	8.922	3.912	3.912	1 LEU	QD2 <==>	9 PRO	HA
5.50	12.012	11.731	6.231	6.231	1 LEU	QD2 <==>	9 PRO	HD3
5.50	12.988	12.440	6.940	6.940	1 LEU	QD1 <==>	9 PRO	HD3
4.47	10.943	10.173	5.703	5.703	1 LEU	QD1 <==>	9 PRO	HB3
4.47	9.744	9.336	4.866	4.866	1 LEU	QD2 <==>	9 PRO	HB3
4.47	11.911	11.225	6.755	6.755	1 LEU	QD1 <==>	9 PRO	HB2
4.47	10.644	10.284	5.814	5.814	1 LEU	QD2 <==>	9 PRO	HB2
4.08	3.480	3.397	-0.683	0.000	8 VAL	H <==>	8 VAL	QG1
3.23	3.433	3.258	0.028	0.028	8 VAL	HA <==>	8 VAL	QG2
3.23	2.438	2.415	-0.815	0.000	8 VAL	HA <==>	8 VAL	QG1
5.15	3.735	3.702	-1.448	0.000	10 GLN	HB2 <==>	11 VAL	H
5.50	12.487	12.373	6.873	6.873	4 TYR	QD <==>	9 PRO	HG2
5.14	14.875	14.558	9.418	9.418	4 TYR	QE <==>	9 PRO	HG2
5.20	2.586	2.520	-2.680	0.000	4 TYR	QD <==>	13 THR	QG2
4.23	8.684	8.523	4.293	4.293	5 THR	QG2 <==>	14 PRO	HD3
3.05	4.950	4.710	1.660	1.660	5 THR	QG2 <==>	13 THR	QG2
5.34	8.672	8.644	3.304	3.304	1 LEU	HA <==>	3 ARG	QB
2.24	2.392	2.387	0.147	0.147	3 ARG	QB <==>	3 ARG	QG
3.03	2.443	2.429	-0.601	0.000	3 ARG	QB <==>	3 ARG	QD
4.97	3.087	3.002	-1.968	0.000	3 ARG	QB <==>	4 TYR	QD
5.34	5.129	4.726	-0.614	0.000	3 ARG	QG <==>	4 TYR	QD
5.34	11.502	11.298	5.958	5.958	4 TYR	QD <==>	9 PRO	QB
4.33	8.006	7.951	3.621	3.621	5 THR	H <==>	14 PRO	QD
3.76	10.251	10.121	6.361	6.361	5 THR	HB <==>	14 PRO	QD
5.35	9.014	8.725	3.375	3.375	5 THR	QG2 <==>	14 PRO	QG
2.73	2.488	2.465	-0.265	0.000	6 LYS	QB <==>	6 LYS	QD
3.64	3.120	2.742	-0.898	0.000	6 LYS	QB <==>	6 LYS	QE
2.74	2.485	2.464	-0.276	0.000	7 LYS	QB <==>	7 LYS	QD
2.53	3.047	2.694	0.164	0.164	7 LYS	QB <==>	7 LYS	QE
5.32	4.485	4.466	-0.854	0.000	9 PRO	QB <==>	10 GLN	HA
2.63	2.468	2.460	-0.170	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.813	4.802	-0.538	0.000	14 PRO	QB <==>	15 THR	HA

```

=====
SUMMARY: 105 unique NOE restraints displayed.
Violations (>0.001Å): 37 (35.2%)
Average violation (for violated restraints): 4.134 Å
Maximum violation: 9.418 Å
Violation severity distribution:
  0-0.5Å: 6 restraints
  >0.5Å: 5 restraints
  >1.0Å: 1 restraints
  >3.0Å: 25 restraints

```

9. DPCA09 (NINTH CLUSTER) :

```

Starting NOE processing using 'carma.distances' at 2025-11-13 01:59:25
Defined 105 unique restraint groups from 257 calculated distances.
Processing 106,973 frames from 'carma.distances' in memory...
Progress: 106973/106973 (100.0%) | Elapsed: 9.8s | Remaining: 0.0s
Finished processing 106,973 frames.

```

NOE Analysis Results:

EXP	<r^-3>	<r^-6>	Deviation	Violation	Atoms Involved			
5.14	5.470	5.141	0.001	0.001	10 GLN	HB2 <==>	12 SER	H
5.14	5.596	5.330	0.190	0.190	10 GLN	HB3 <==>	12 SER	H
5.41	3.764	3.504	-1.906	0.000	11 VAL	HB <==>	12 SER	H
5.50	3.046	2.874	-2.626	0.000	10 GLN	H <==>	10 GLN	QG
5.50	5.537	5.338	-0.162	0.000	10 GLN	QG <==>	12 SER	H
5.50	8.160	7.218	1.718	1.718	5 THR	H <==>	13 THR	QG2
5.50	2.791	2.687	-2.813	0.000	13 THR	H <==>	13 THR	QG2
4.92	3.632	3.319	-1.601	0.000	15 THR	H <==>	15 THR	QG2
5.50	2.501	2.415	-3.085	0.000	7 LYS	H <==>	7 LYS	QG
5.50	3.765	3.503	-1.997	0.000	6 LYS	QG <==>	7 LYS	H
5.50	3.998	3.779	-1.721	0.000	7 LYS	H <==>	7 LYS	QD
5.50	2.665	2.582	-2.918	0.000	6 LYS	H <==>	6 LYS	QG
5.50	4.212	4.031	-1.469	0.000	6 LYS	H <==>	6 LYS	QD
5.50	4.760	4.566	-0.934	0.000	6 LYS	H <==>	6 LYS	QE
3.75	4.726	4.338	0.588	0.588	7 LYS	HA <==>	7 LYS	QE
3.08	4.853	4.614	1.534	1.534	6 LYS	HA <==>	6 LYS	QE
2.97	2.484	2.474	-0.496	0.000	6 LYS	HB2 <==>	6 LYS	QG
2.88	2.463	2.451	-0.429	0.000	7 LYS	HB2 <==>	7 LYS	QG
5.50	2.894	2.682	-2.818	0.000	3 ARG	H <==>	3 ARG	HG2
5.50	3.259	3.090	-2.410	0.000	3 ARG	H <==>	3 ARG	HG3
4.33	4.142	3.691	-0.639	0.000	3 ARG	HA <==>	3 ARG	QD
3.86	2.644	2.528	-1.332	0.000	7 LYS	HB3 <==>	8 VAL	H
4.08	2.997	2.878	-1.202	0.000	8 VAL	H <==>	8 VAL	QG2
5.15	3.829	3.649	-1.501	0.000	10 GLN	HB3 <==>	11 VAL	H
3.11	3.428	3.343	0.233	0.233	9 PRO	HA <==>	10 GLN	H
3.18	2.349	2.331	-0.849	0.000	14 PRO	HA <==>	15 THR	H
3.36	2.396	2.364	-0.996	0.000	12 SER	HA <==>	13 THR	H
5.50	3.497	3.386	-2.114	0.000	8 VAL	H <==>	9 PRO	HD2
5.50	3.014	2.893	-2.607	0.000	5 THR	H <==>	5 THR	QG2
2.92	2.251	2.230	-0.690	0.000	15 THR	HA <==>	16 LEU	H
2.92	2.558	2.538	-0.382	0.000	13 THR	HA <==>	13 THR	HB
5.50	2.683	2.559	-2.941	0.000	2 VAL	HB <==>	3 ARG	H
5.50	4.369	4.232	-1.268	0.000	4 TYR	H <==>	5 THR	H
2.86	3.805	3.757	0.897	0.897	8 VAL	H <==>	8 VAL	HB
4.28	2.908	2.599	-1.681	0.000	1 LEU	HA <==>	1 LEU	QD1
4.28	2.597	2.413	-1.867	0.000	1 LEU	HA <==>	1 LEU	QD2
4.65	7.427	5.766	1.116	1.116	1 LEU	QD1 <==>	4 TYR	QD
4.65	7.759	6.060	1.410	1.410	1 LEU	QD2 <==>	4 TYR	QD
5.50	7.026	5.181	-0.319	0.000	1 LEU	QD1 <==>	4 TYR	QE
5.50	7.460	5.597	0.097	0.097	1 LEU	QD2 <==>	4 TYR	QE
4.77	6.449	5.361	0.591	0.591	3 ARG	HB3 <==>	4 TYR	QE
3.65	3.204	3.048	-0.602	0.000	3 ARG	HA <==>	3 ARG	HG3
3.65	3.101	2.936	-0.714	0.000	3 ARG	HA <==>	3 ARG	HG2
4.77	6.830	6.109	1.339	1.339	3 ARG	HB2 <==>	4 TYR	QE
5.29	3.579	3.430	-1.860	0.000	3 ARG	HA <==>	4 TYR	QD
3.17	3.004	2.936	-0.234	0.000	4 TYR	HA <==>	4 TYR	QD
5.11	5.213	5.041	-0.069	0.000	4 TYR	QD <==>	9 PRO	HD3
5.50	5.959	5.711	0.211	0.211	4 TYR	QE <==>	9 PRO	HD3
2.84	2.460	2.437	-0.403	0.000	13 THR	HA <==>	13 THR	QG2
2.40	2.375	2.369	-0.031	0.000	7 LYS	QG <==>	7 LYS	QD
2.88	2.462	2.450	-0.430	0.000	7 LYS	HB3 <==>	7 LYS	QG
3.82	2.431	2.407	-1.413	0.000	7 LYS	QG <==>	7 LYS	QE
3.79	3.536	3.080	-0.710	0.000	7 LYS	HA <==>	7 LYS	QD
3.63	2.657	2.585	-1.045	0.000	7 LYS	HA <==>	7 LYS	QG
2.97	2.454	2.444	-0.526	0.000	6 LYS	HB3 <==>	6 LYS	QG
2.40	2.375	2.369	-0.031	0.000	6 LYS	QG <==>	6 LYS	QD
3.63	3.823	3.357	-0.273	0.000	6 LYS	HA <==>	6 LYS	QD
3.45	2.576	2.535	-0.915	0.000	6 LYS	HA <==>	6 LYS	QG
5.50	5.568	5.506	0.006	0.006	7 LYS	HA <==>	8 VAL	HB
3.86	2.874	2.739	-1.121	0.000	10 GLN	HA <==>	10 GLN	QG

3.68	3.493	3.310	-0.370	0.000	11 VAL	H <==>	11 VAL	HB
4.56	2.603	2.521	-2.039	0.000	12 SER	QB <==>	13 THR	H
5.50	4.560	4.367	-1.133	0.000	7 LYS	QD <==>	8 VAL	H
3.86	3.358	3.142	-0.718	0.000	7 LYS	HB2 <==>	8 VAL	H
2.85	2.307	2.288	-0.562	0.000	7 LYS	HA <==>	8 VAL	H
5.50	6.288	6.184	0.684	0.684	14 PRO	HA <==>	16 LEU	H
2.40	2.628	2.551	0.151	0.151	8 VAL	HA <==>	9 PRO	HD3
2.41	2.410	2.369	-0.041	0.000	8 VAL	HA <==>	9 PRO	HD2
3.46	10.522	10.169	6.709	6.709	5 THR	HA <==>	14 PRO	HD2
3.46	10.247	9.995	6.535	6.535	5 THR	HA <==>	14 PRO	HD3
4.23	8.860	7.916	3.686	3.686	5 THR	QG2 <==>	14 PRO	HD2
4.11	4.116	3.501	-0.609	0.000	8 VAL	HB <==>	9 PRO	HD2
5.01	12.346	9.977	4.967	4.967	1 LEU	QD1 <==>	9 PRO	HA
5.01	12.328	9.196	4.186	4.186	1 LEU	QD2 <==>	9 PRO	HA
5.50	13.114	10.387	4.887	4.887	1 LEU	QD2 <==>	9 PRO	HD3
5.50	12.939	9.979	4.479	4.479	1 LEU	QD1 <==>	9 PRO	HD3
4.47	10.519	7.593	3.123	3.123	1 LEU	QD1 <==>	9 PRO	HB3
4.47	10.739	7.656	3.186	3.186	1 LEU	QD2 <==>	9 PRO	HB3
4.47	10.924	8.020	3.550	3.550	1 LEU	QD1 <==>	9 PRO	HB2
4.47	11.117	8.325	3.855	3.855	1 LEU	QD2 <==>	9 PRO	HB2
4.08	3.010	2.837	-1.243	0.000	8 VAL	H <==>	8 VAL	QG1
3.23	3.076	2.915	-0.315	0.000	8 VAL	HA <==>	8 VAL	QG2
3.23	2.384	2.364	-0.866	0.000	8 VAL	HA <==>	8 VAL	QG1
5.15	3.671	3.529	-1.621	0.000	10 GLN	HB2 <==>	11 VAL	H
5.50	5.134	4.898	-0.602	0.000	4 TYR	QD <==>	9 PRO	HG2
5.14	5.522	5.226	0.086	0.086	4 TYR	QE <==>	9 PRO	HG2
5.20	7.064	6.540	1.340	1.340	4 TYR	QD <==>	13 THR	QG2
4.23	9.127	8.441	4.211	4.211	5 THR	QG2 <==>	14 PRO	HD3
3.05	7.727	5.405	2.355	2.355	5 THR	QG2 <==>	13 THR	QG2
5.34	7.201	7.031	1.691	1.691	1 LEU	HA <==>	3 ARG	QB
2.24	2.379	2.372	0.132	0.132	3 ARG	QB <==>	3 ARG	QG
3.03	2.457	2.437	-0.593	0.000	3 ARG	QB <==>	3 ARG	QD
4.97	4.965	4.474	-0.496	0.000	3 ARG	QB <==>	4 TYR	QD
5.34	5.587	5.064	-0.276	0.000	3 ARG	QG <==>	4 TYR	QD
5.34	3.472	3.274	-2.066	0.000	4 TYR	QD <==>	9 PRO	QB
4.33	7.400	7.179	2.849	2.849	5 THR	H <==>	14 PRO	QD
3.76	10.585	10.202	6.442	6.442	5 THR	HB <==>	14 PRO	QD
5.35	8.814	8.362	3.012	3.012	5 THR	QG2 <==>	14 PRO	QG
2.73	2.458	2.442	-0.288	0.000	6 LYS	QB <==>	6 LYS	QD
3.64	3.203	2.804	-0.836	0.000	6 LYS	QB <==>	6 LYS	QE
2.74	2.475	2.457	-0.283	0.000	7 LYS	QB <==>	7 LYS	QD
2.53	3.158	2.769	0.239	0.239	7 LYS	QB <==>	7 LYS	QE
5.32	4.336	4.311	-1.009	0.000	9 PRO	QB <==>	10 GLN	HA
2.63	2.411	2.400	-0.230	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.488	4.463	-0.877	0.000	14 PRO	QB <==>	15 THR	HA

=====

SUMMARY: 105 unique NOE restraints displayed.

Violations (>0.001Å): 36 (34.3%)

Average violation (for violated restraints): 2.286 Å

Maximum violation: 6.709 Å

Violation severity distribution:

0-0.5Å: 9 restraints

>0.5Å: 4 restraints

>1.0Å: 7 restraints

>2.0Å: 2 restraints

>3.0Å: 14 restraints

10. DPCA10 (TENTH CLUSTER):

Starting NOE processing using 'carma.distances' at 2025-11-13 02:00:07
 Defined 105 unique restraint groups from 257 calculated distances.
 Processing 37,756 frames from 'carma.distances' in memory...
 Progress: 37756/37756 (100.0%) | Elapsed: 3.4s | Remaining: 0.0s
 Finished processing 37,756 frames.

NOE Analysis Results:

EXP	<r ⁻³ >	<r ⁻⁶ >	Deviation	Violation	Atoms Involved			
5.14	6.413	6.363	1.223	1.223	10 GLN	HB2 <==>	12 SER	H
5.14	6.041	5.935	0.795	0.795	10 GLN	HB3 <==>	12 SER	H
5.41	2.537	2.470	-2.940	0.000	11 VAL	HB <==>	12 SER	H
5.50	3.026	2.937	-2.563	0.000	10 GLN	H <==>	10 GLN	QG
5.50	7.376	6.941	1.441	1.441	10 GLN	QG <==>	12 SER	H
5.50	4.063	3.902	-1.598	0.000	5 THR	H <==>	13 THR	QG2
5.50	2.578	2.515	-2.985	0.000	13 THR	H <==>	13 THR	QG2
4.92	2.859	2.782	-2.138	0.000	15 THR	H <==>	15 THR	QG2
5.50	2.920	2.720	-2.780	0.000	7 LYS	H <==>	7 LYS	QG
5.50	4.051	3.956	-1.544	0.000	6 LYS	QG <==>	7 LYS	H
5.50	4.274	4.040	-1.460	0.000	7 LYS	H <==>	7 LYS	QD
5.50	2.898	2.832	-2.668	0.000	6 LYS	H <==>	6 LYS	QG
5.50	4.424	4.327	-1.173	0.000	6 LYS	H <==>	6 LYS	QD
5.50	5.178	5.079	-0.421	0.000	6 LYS	H <==>	6 LYS	QE
3.75	4.472	3.977	0.227	0.227	7 LYS	HA <==>	7 LYS	QE
3.08	4.820	4.590	1.510	1.510	6 LYS	HA <==>	6 LYS	QE
2.97	2.458	2.448	-0.522	0.000	6 LYS	HB2 <==>	6 LYS	QG
2.88	2.460	2.449	-0.431	0.000	7 LYS	HB2 <==>	7 LYS	QG
5.50	2.884	2.716	-2.784	0.000	3 ARG	H <==>	3 ARG	HG2
5.50	3.411	3.181	-2.319	0.000	3 ARG	H <==>	3 ARG	HG3
4.33	2.770	2.553	-1.777	0.000	3 ARG	HA <==>	3 ARG	QD
3.86	3.334	3.175	-0.685	0.000	7 LYS	HB3 <==>	8 VAL	H
4.08	2.185	2.147	-1.933	0.000	8 VAL	H <==>	8 VAL	QG2
5.15	2.690	2.603	-2.547	0.000	10 GLN	HB3 <==>	11 VAL	H
3.11	2.243	2.229	-0.881	0.000	9 PRO	HA <==>	10 GLN	H
3.18	2.276	2.257	-0.923	0.000	14 PRO	HA <==>	15 THR	H
3.36	2.842	2.806	-0.554	0.000	12 SER	HA <==>	13 THR	H
5.50	5.162	4.979	-0.521	0.000	8 VAL	H <==>	9 PRO	HD2
5.50	3.137	3.077	-2.423	0.000	5 THR	H <==>	5 THR	QG2
2.92	3.302	3.244	0.324	0.324	15 THR	HA <==>	16 LEU	H
2.92	2.481	2.468	-0.452	0.000	13 THR	HA <==>	13 THR	HB
5.50	3.274	2.988	-2.512	0.000	2 VAL	HB <==>	3 ARG	H
5.50	4.298	4.072	-1.428	0.000	4 TYR	H <==>	5 THR	H
2.86	3.620	3.547	0.687	0.687	8 VAL	H <==>	8 VAL	HB
4.28	2.760	2.521	-1.759	0.000	1 LEU	HA <==>	1 LEU	QD1
4.28	2.729	2.487	-1.793	0.000	1 LEU	HA <==>	1 LEU	QD2
4.65	6.955	5.989	1.339	1.339	1 LEU	QD1 <==>	4 TYR	QD
4.65	7.353	6.547	1.897	1.897	1 LEU	QD2 <==>	4 TYR	QD
5.50	6.269	5.068	-0.432	0.000	1 LEU	QD1 <==>	4 TYR	QE
5.50	6.952	5.772	0.272	0.272	1 LEU	QD2 <==>	4 TYR	QE
4.77	7.161	6.796	2.026	2.026	3 ARG	HB3 <==>	4 TYR	QE
3.65	3.219	3.049	-0.601	0.000	3 ARG	HA <==>	3 ARG	HG3
3.65	2.986	2.874	-0.776	0.000	3 ARG	HA <==>	3 ARG	HG2
4.77	7.116	6.993	2.223	2.223	3 ARG	HB2 <==>	4 TYR	QE
5.29	3.949	3.794	-1.496	0.000	3 ARG	HA <==>	4 TYR	QD
3.17	2.546	2.491	-0.679	0.000	4 TYR	HA <==>	4 TYR	QD
5.11	7.778	7.116	2.006	2.006	4 TYR	QD <==>	9 PRO	HD3
5.50	7.414	6.038	0.538	0.538	4 TYR	QE <==>	9 PRO	HD3
2.84	2.480	2.459	-0.381	0.000	13 THR	HA <==>	13 THR	QG2
2.40	2.374	2.368	-0.032	0.000	7 LYS	QG <==>	7 LYS	QD
2.88	2.454	2.443	-0.437	0.000	7 LYS	HB3 <==>	7 LYS	QG
3.82	2.446	2.418	-1.402	0.000	7 LYS	QG <==>	7 LYS	QE
3.79	3.321	2.919	-0.871	0.000	7 LYS	HA <==>	7 LYS	QD

3.63	2.595	2.545	-1.085	0.000	7 LYS	HA <==>	7 LYS	QG
2.97	2.469	2.460	-0.510	0.000	6 LYS	HB3 <==>	6 LYS	QG
2.40	2.381	2.375	-0.025	0.000	6 LYS	QG <==>	6 LYS	QD
3.63	3.519	3.075	-0.555	0.000	6 LYS	HA <==>	6 LYS	QD
3.45	2.605	2.561	-0.889	0.000	6 LYS	HA <==>	6 LYS	QG
5.50	6.531	6.518	1.018	1.018	7 LYS	HA <==>	8 VAL	HB
3.86	2.645	2.601	-1.259	0.000	10 GLN	HA <==>	10 GLN	QG
3.68	3.621	3.484	-0.196	0.000	11 VAL	H <==>	11 VAL	HB
4.56	4.160	4.066	-0.494	0.000	12 SER	QB <==>	13 THR	H
5.50	5.384	5.297	-0.203	0.000	7 LYS	QD <==>	8 VAL	H
3.86	3.342	3.198	-0.662	0.000	7 LYS	HB2 <==>	8 VAL	H
2.85	3.423	3.397	0.547	0.547	7 LYS	HA <==>	8 VAL	H
5.50	4.140	4.101	-1.399	0.000	14 PRO	HA <==>	16 LEU	H
2.40	2.230	2.201	-0.199	0.000	8 VAL	HA <==>	9 PRO	HD3
2.41	2.766	2.692	0.282	0.282	8 VAL	HA <==>	9 PRO	HD2
3.46	7.749	7.546	4.086	4.086	5 THR	HA <==>	14 PRO	HD2
3.46	8.743	8.497	5.037	5.037	5 THR	HA <==>	14 PRO	HD3
4.23	4.863	4.384	0.154	0.154	5 THR	QG2 <==>	14 PRO	HD2
4.11	2.347	2.310	-1.800	0.000	8 VAL	HB <==>	9 PRO	HD2
5.01	13.256	11.220	6.210	6.210	1 LEU	QD1 <==>	9 PRO	HA
5.01	13.688	11.840	6.830	6.830	1 LEU	QD2 <==>	9 PRO	HA
5.50	15.963	10.050	4.550	4.550	1 LEU	QD2 <==>	9 PRO	HD3
5.50	15.624	11.056	5.556	5.556	1 LEU	QD1 <==>	9 PRO	HD3
4.47	13.712	10.410	5.940	5.940	1 LEU	QD1 <==>	9 PRO	HB3
4.47	14.125	10.079	5.609	5.609	1 LEU	QD2 <==>	9 PRO	HB3
4.47	14.724	9.367	4.897	4.897	1 LEU	QD1 <==>	9 PRO	HB2
4.47	15.235	10.240	5.770	5.770	1 LEU	QD2 <==>	9 PRO	HB2
4.08	2.625	2.555	-1.525	0.000	8 VAL	H <==>	8 VAL	QG1
3.23	3.665	3.581	0.351	0.351	8 VAL	HA <==>	8 VAL	QG2
3.23	2.431	2.413	-0.817	0.000	8 VAL	HA <==>	8 VAL	QG1
5.15	3.467	3.255	-1.895	0.000	10 GLN	HB2 <==>	11 VAL	H
5.50	8.480	7.909	2.409	2.409	4 TYR	QD <==>	9 PRO	HG2
5.14	7.649	6.964	1.824	1.824	4 TYR	QE <==>	9 PRO	HG2
5.20	2.986	2.813	-2.387	0.000	4 TYR	QD <==>	13 THR	QG2
4.23	5.660	5.118	0.888	0.888	5 THR	QG2 <==>	14 PRO	HD3
3.05	4.739	4.324	1.274	1.274	5 THR	QG2 <==>	13 THR	QG2
5.34	7.855	7.769	2.429	2.429	1 LEU	HA <==>	3 ARG	QB
2.24	2.362	2.356	0.116	0.116	3 ARG	QB <==>	3 ARG	QG
3.03	2.532	2.508	-0.522	0.000	3 ARG	QB <==>	3 ARG	QD
4.97	5.674	5.470	0.500	0.500	3 ARG	QB <==>	4 TYR	QD
5.34	6.290	5.760	0.420	0.420	3 ARG	QG <==>	4 TYR	QD
5.34	6.403	5.729	0.389	0.389	4 TYR	QD <==>	9 PRO	QB
4.33	6.318	6.063	1.733	1.733	5 THR	H <==>	14 PRO	QD
3.76	7.696	7.397	3.637	3.637	5 THR	HB <==>	14 PRO	QD
5.35	5.371	4.496	-0.854	0.000	5 THR	QG2 <==>	14 PRO	QG
2.73	2.458	2.443	-0.287	0.000	6 LYS	QB <==>	6 LYS	QD
3.64	3.514	3.059	-0.581	0.000	6 LYS	QB <==>	6 LYS	QE
2.74	2.492	2.471	-0.269	0.000	7 LYS	QB <==>	7 LYS	QD
2.53	3.081	2.716	0.186	0.186	7 LYS	QB <==>	7 LYS	QE
5.32	4.546	4.529	-0.791	0.000	9 PRO	QB <==>	10 GLN	HA
2.63	2.450	2.440	-0.190	0.000	10 GLN	HA <==>	10 GLN	QB
5.34	4.724	4.706	-0.634	0.000	14 PRO	QB <==>	15 THR	HA

=====

SUMMARY: 105 unique NOE restraints displayed.

Violations (>0.001Å): 41 (39.0%)

Average violation (for violated restraints): 2.174 Å

Maximum violation: 6.830 Å

Violation severity distribution:

0-0.5Å: 11 restraints

>0.5Å: 5 restraints

>1.0Å: 9 restraints

>2.0Å: 5 restraints

Chapter 7: Conclusion and Final Remarks

7.1 Summary of Contributions

The NOEmd program, as it's presented in this dissertation, is a concise and rapid, automated computational tool for the validation of molecular dynamics (MD) trajectories against nuclear overhauser effect (NOE) distance restraints obtained experimentally. By merging structural bioinformatics tools with solution-state NMR data, NOEmd addresses a number of the traditional limitations associated with the refinement process in structural biology. The primary technical and methodological innovations of this work are summarized as follows:

1. **IUPAC Naming Automization:** A major cornerstone of the NOEmd is to automate standard file-level parsing of topology (PSF) and restraint (MR) file types. Using an active translation mechanism based on established nomenclature and utilizing omni-directional forward and reverse dictionaries, it links pseudo atom labels to atomic structure names explicitly defined in memory, thus avoiding human error and guaranteeing data consistency on a core level.
2. **Unified Degeneracy Resolution:** One of the key features of this architecture is the use of an extremely efficient chemical and structure proton degeneracy resolution strategy. This program maps individual degenerate atom-pair columns from an internal, structure tracking dictionary. Using a single pass algorithm, trajectory segments are evaluated in memory to reduce the multi-column matrix into a single minimum distance per restraint group per frame that provides the correct physical basis for the actual ensemble averaging.
3. Precise analytical validation is also performed by having the analytical core execute two consecutive numerical integrations that compute, in one pass, the inverse third and inverse sixth time-ensemble averages using the raw trajectory data file. The `run_noe_averaging` function compares these values to experimental values and returns an exact deviations array. This is compiled into a properly tabulated, color coded report of summary global violations distributions.
4. **High-Performance command-line Integration:** By unifying data aggregation, structural indexing, sub-process conjunction with the external carma passthrough, data matrix down-sampling, and physical error quantification into a singular cohesive script, NOEmd transforms a historically cumbersome multi-step workload into a rapid

command-line function. It greatly increases calculation reproducibility, reducing hours of manual manipulation into seconds of processing time.

7.2 Final Remarks

Unraveling and cross-validating the three-dimensional structures of biomolecules remains an indispensable and fundamental part of modern molecular biology and structure-based drug design. NOEmd provides an advanced structural computational biology asset that automates the laborious tasks of multi-source data merging and physical degeneracy grouping. By combining meticulous nomenclature dictionary translation with rapid in-memory processing and clear quantitative error auditing, this work makes the critical task of validating theoretical molecular trajectories against real-world solution-state NMR data vastly more accessible, efficient, and reproducible.

Source Code

Please note that the version of the source code provided below is constantly subject to change. For versions that are up to date, as well as to keep up with improvements, new implementations and otherwise changes that may not be documented on the snapshot that's attached please visit the Github repository at : <https://github.com/tsevisaggelos/nOeMD/>

```
"""
```

```
"""
```

NOEmd : a computer program to automate the comparison between experimental NMR data and molecular dynamics simulations

This script processes molecular dynamics simulation data against experimental restraints, typically from NMR (NOE) experiments. It performs the following key functions:

1. Parses a STAR/.mr file for distance restraints and a PSF file for system topology.
2. Maps the atom names from the restraint file (e.g., using Markley nomenclature) to the corresponding atom IDs in the PSF file.
3. Handles degenerate restraints (e.g., methyl groups) by identifying all possible atom pairs.
4. Generates an input file for CARMA to calculate distances for these pairs across a DCD trajectory.
5. (Optional) Invokes CARMA to perform the distance calculations.
6. Processes the output from CARMA, resolving degeneracies by selecting the minimum distance for each restraint group in each frame.
7. Calculates $\langle r^{-3} \rangle$ and $\langle r^{-6} \rangle$ time averages for each restraint directly in memory.
8. Generates a final report comparing the experimental distances to the

```

        simulation-derived averages and calculates violations.
"""

import re
import sys
import os
import subprocess
import time
import math
import numpy as np
from collections import defaultdict
import mmap
import datetime
import concurrent.futures
from multiprocessing import cpu_count

# Color codes
class Colors:
    HEADER = '\033[95m'
    OKBLUE = '\033[94m'
    OKCYAN = '\033[96m'
    OKGREEN = '\033[92m'
    WARNING = '\033[93m'
    FAIL = '\033[91m'
    ENDC = '\033[0m'
    BOLD = '\033[1m'
    UNDERLINE = '\033[4m'

# Amino acid specific mappings from Markley nomenclature

```

```

markley_map = {
  "ALA": {"QB": ["HB1", "HB2", "HB3"]},
  "VAL": {
    "QG1": ["HG11", "HG12", "HG13"],
    "QG2": ["HG21", "HG22", "HG23"]
  },
  "LEU": {
    "QD1": ["HD11", "HD12", "HD13"],
    "QD2": ["HD21", "HD22", "HD23"]
  },
  "ILE": {
    "QG1": ["HG12", "HG13"],
    "QG2": ["HG21", "HG22", "HG23"],
    "QD1": ["HD11", "HD12", "HD13"]
  },
  "THR": {
    "QG1": ["HG1"],
    "HG2": "HG2",
    "QG2": ["HG21", "HG22", "HG23"],
    "HG3": "HG3"
  },
  "SER": {
    "QB": ["HB2", "HB3"],
    "HG": "HG"
  },
  "CYS": {
    "QB": ["HB2", "HB3"],
    "HG": "HG"
  },
}

```

```
"MET": {
  "QB": ["HB2", "HB3"],
  "QG": ["HG2", "HG3"],
  "QE": ["HE1", "HE2", "HE3"]
},
"PHE": {
  "QB": ["HB2", "HB3"],
  "QD": ["HD1", "HD2"],
  "QE": ["HE1", "HE2"],
  "HZ": "HZ"
},
"TYR": {
  "QB": ["HB2", "HB3"],
  "QD": ["HD1", "HD2"],
  "QE": ["HE1", "HE2"],
  "HH": "HH"
},
"TRP": {
  "QB": ["HB2", "HB3"],
  "QD1": "HD1",
  "QE3": "HE3",
  "QZ2": "HZ2",
  "QZ3": "HZ3",
  "QH2": "HH2"
},
"GLN": {
  "QB": ["HB2", "HB3"],
  "QG": ["HG2", "HG3"],
  "QE": ["HE21", "HE22"]
}
```

```

},
"GLU": {
    "QB": ["HB2", "HB3"],
    "QG": ["HG2", "HG3"],
    "QD": ["HD21", "HD22"]
},
"ASN": {
    "QB": ["HB2", "HB3"],
    "QD": ["HD21", "HD22"]
},
"ASP": {
    "QB": ["HB2", "HB3"],
    "QD": ["HD21", "HD22"]
},
"LYS": {
    "QB": ["HB2", "HB3"],
    "QG": ["HG2", "HG3"],
    "QD": ["HD2", "HD3"],
    "QE": ["HE2", "HE3"],
    "QZ": ["HZ1", "HZ2", "HZ3"]
},
"ARG": {
    "QB": ["HB2", "HB3"],
    "QG": ["HG2", "HG3"],
    "QD": ["HD2", "HD3"],
    "QH": ["HH11", "HH12", "HH21", "HH22"]
},
"HIS": {
    "QB": ["HB2", "HB3"],

```

```

        "QD": ["HD1"],
        "QE": ["HE1"]
    },
    "PRO": {
        "QB": ["HB2", "HB3"],
        "QG": ["HG2", "HG3"],
        "QD": ["HD2", "HD3"]
    }
}

# Generate reverse mappings
reverse_markley_map = {}
for res_name, mappings in markley_map.items():
    res_reverse_map = {}
    for k, v in mappings.items():
        if isinstance(v, list):
            for item in v:
                res_reverse_map[item] = k
        else:
            res_reverse_map[v] = k
    reverse_markley_map[res_name] = res_reverse_map

def parse_mr_file(mr_file_path):
    """Returns atom pairs AND their experimental distances."""
    entries = []
    entry_pattern = re.compile(
        r"(\d+)\s+(\w+)\s+(\w+)\s+(\d+)\s+(\w+)\s+(\w+)\s+(\w+)\s+(\[d.]+\)"
    )
    with open(mr_file_path, 'r') as mr_file:

```

```

for line in mr_file:
    if line.startswith('*') or not line.strip():
        continue
    match = entry_pattern.search(line)
    if match:
        atom1 = (match.group(1), match.group(2), match.group(3))
        atom2 = (match.group(4), match.group(5), match.group(6))
        distance = float(match.group(7))
        entries.append((atom1, atom2, distance))
return entries

def find_all_atom_entries(psf_file_path, residue_number, residue_name, atom_type):
    """Optimized PSF file parser with precompiled regex"""
    entries = []
    atom_types_to_try = {atom_type}

    if residue_name in markley_map and atom_type in markley_map[residue_name]:
        mapped = markley_map[residue_name][atom_type]
        if isinstance(mapped, list):
            atom_types_to_try.update(mapped)
        else:
            atom_types_to_try.add(mapped)

    if residue_name in reverse_markley_map and atom_type in reverse_markley_map[residue_name]:
        atom_types_to_try.add(reverse_markley_map[residue_name][atom_type])

    pattern =
re.compile(rf"^\s*(\d+)\s+\w\s+{residue_number}\s+{residue_name}\s+({'|'.join(atom_types_to_try)})\s")

```

```

with open(psf_file_path, 'r') as psf_file:
    for line in psf_file:
        match = pattern.match(line)
        if match:
            parts = line.strip().split()
            parts[4] = atom_type # Keep original atom name
            entries.append(" ".join(parts))

    return entries if entries else [f"???? ? {residue_number:>4} {residue_name:>3} {atom_type:>4}"]

def find_atom_ids_for_pairs(mr_file_path, psf_file_path):
    """Returns formatted pairs with their experimental distances."""
    entries = parse_mr_file(mr_file_path)
    formatted_output = []
    seen_pairs = set()

    # Build cache for all unique atoms
    unique_atoms = set()
    for atom1, atom2, _ in entries:
        unique_atoms.add(atom1)
        unique_atoms.add(atom2)

    atom_entries_cache = {}
    for atom in unique_atoms:
        atom_entries = find_all_atom_entries(psf_file_path, *atom)
        if not any("?????" in entry for entry in atom_entries):
            atom_entries_cache[atom] = atom_entries

    # Process pairs with distances

```

```

for atom1, atom2, exp_distance in entries:
    if atom1 not in atom_entries_cache or atom2 not in atom_entries_cache:
        continue

    atom1_entries = atom_entries_cache[atom1]
    atom2_entries = atom_entries_cache[atom2]

    for a1_entry in atom1_entries:
        for a2_entry in atom2_entries:
            atom1_id = a1_entry.split()[0]
            atom2_id = a2_entry.split()[0]

            pair_id = tuple(sorted((atom1_id, atom2_id)))
            if pair_id in seen_pairs:
                continue
            seen_pairs.add(pair_id)

            a1_parts = a1_entry.split()
            a2_parts = a2_entry.split()

            is_degenerate = (len(atom1_entries) > 1 or len(atom2_entries) > 1)

            formatted_line = (
                f"{atom1_id:>4} {atom2_id:>4} {exp_distance:6.2f} *** "
                f"{a1_parts[0]:>4} {a1_parts[1]:<1} {a1_parts[2]:>4} {a1_parts[3]:>3} {atom1[2]:>4} <==> "
                f"{a2_parts[0]:>4} {a2_parts[1]:<1} {a2_parts[2]:>4} {a2_parts[3]:>3} {atom2[2]:>4}"
                f"{' (Degenerate)' if is_degenerate else ''}"
            )
            formatted_output.append(formatted_line)

```

```

    return formatted_output, len(formatted_output) == 0

def calculate_noe_sums_in_memory(formatted_atom_pairs, carma_file='carma.distances'):
    """
    Reads the raw carma.distances file, resolves degeneracies for each frame
    in memory, and directly calculates the sum of r^-3 and r^-6 without
    writing an intermediate file.
    """
    if not os.path.exists(carma_file):
        print(f"{Colors.FAIL}ERROR: {carma_file} not found for NOE processing.{Colors.ENDC}")
        return None

    # Step 1: Group columns by their restraint definition. This tells us which
    # columns in the carma.distances file belong to the same restraint.
    # We also create a list of the unique restraint keys in the order they appear.
    column_groups = defaultdict(list)
    unique_restraint_keys = []
    # This map will link the index of a column in the raw carma file to the
    # index in our unique_restraint_keys list.
    raw_col_to_unique_group_idx = {}

    temp_unique_id_pairs_for_carma = set()
    col_idx_counter = 0
    for pair_desc in formatted_atom_pairs:
        parts = pair_desc.split("****")
        id_parts = parts[0].strip().split()
        atom_id1, atom_id2 = id_parts[0], id_parts[1]

```

```

# This logic must match how carmainput_for_distcalc.txt is created
try:
    id_pair = tuple(sorted((int(atom_id1), int(atom_id2))))
    if id_pair in temp_unique_id_pairs_for_carma:
        continue # This atom ID pair is a duplicate, skip
    temp_unique_id_pairs_for_carma.add(id_pair)
except (ValueError, IndexError):
    continue # Skip error/malformed lines

# Define the group key using residue/atom names
detail_parts = parts[1].split("<==>")
left_psf = detail_parts[0].strip().split()
right_psf_part = detail_parts[1].strip()
right_psf = right_psf_part.split(" (Degenerate)")[0].strip().split()

if len(left_psf) < 5 or len(right_psf) < 5:
    continue

k1 = (left_psf[2], left_psf[3], left_psf[4])
k2 = (right_psf[2], right_psf[3], right_psf[4])
group_key = tuple(sorted((k1, k2)))

if group_key not in column_groups:
    unique_restraint_keys.append(group_key)

# The current raw column index corresponds to this group
group_idx = unique_restraint_keys.index(group_key)
column_groups[group_key].append(col_idx_counter)
raw_col_to_unique_group_idx[col_idx_counter] = group_idx

```

```

        col_idx_counter += 1

num_raw_cols = col_idx_counter
num_unique_groups = len(unique_restraint_keys)

print(f"{Colors.OKCYAN}Defined {num_unique_groups} unique restraint groups from {num_raw_cols} calculated
distances.{Colors.ENDC}")

# Step 2: Initialize sum arrays and frame counter
sum_r3 = np.zeros(num_unique_groups)
sum_r6 = np.zeros(num_unique_groups)
n_frames = 0
total_lines = sum(1 for line in open(carma_file, 'r') if line.strip())

if total_lines == 0:
    print(f"{Colors.FAIL}ERROR: Distance file '{carma_file}' is empty.{Colors.ENDC}")
    return None

print(f"{Colors.OKBLUE}Processing {total_lines:,} frames from '{carma_file}' in memory...{Colors.ENDC}")
start_time = time.time()

# Step 3: Process the carma.distances file line by line
with open(carma_file, 'r') as f:
    for line_num, line in enumerate(f, 1):
        line_strip = line.strip()
        if not line_strip or line_strip.startswith('#'):
            continue

```

```

parts = line_strip.split()
distances = parts[1:] # parts[0] is frame index

if len(distances) != num_raw_cols:
    print(f"\n{Colors.WARNING}Line {line_num}: Expected {num_raw_cols} distances, found
{len(distances)}. Skipping line.{Colors.ENDC}")
    continue

n_frames += 1

# For each unique group, find the minimum distance among its members for this frame
for i, group_key in enumerate(unique_restraint_keys):
    # Get the raw column indices for this group
    raw_indices_for_group = column_groups[group_key]

    # Extract the distances for this group from the full list of distances
    try:
        group_distances = [float(distances[j]) for j in raw_indices_for_group]
    except (ValueError, IndexError) as e:
        print(f"\n{Colors.WARNING}Error processing distances for group {i} on line {line_num}.
Details: {e}. Skipping group for this frame.{Colors.ENDC}")
        continue

    if not group_distances:
        continue

    # Find the minimum distance for this group in this frame
    min_dist = min(group_distances)

```

```

    # Add to the sums (with an epsilon to prevent division by zero)
    sum_r3[i] += 1.0 / ((min_dist + 1e-9) ** 3)
    sum_r6[i] += 1.0 / ((min_dist + 1e-9) ** 6)

# Progress reporting
if n_frames % 100 == 0 or n_frames == total_lines:
    elapsed = time.time() - start_time
    remaining = (elapsed / n_frames) * (total_lines - n_frames) if n_frames > 0 else 0
    print(
        f"\r{Colors.OKBLUE}Progress: {n_frames}/{total_lines} "
        f"({n_frames/total_lines*100:.1f}%) | "
        f"Elapsed: {elapsed:.1f}s | "
        f"Remaining: {remaining:.1f}s{Colors.ENDC}",
        end="", flush=True
    )
print(f"\n{Colors.OKGREEN}Finished processing {n_frames:,} frames.{Colors.ENDC}")
if n_frames == 0:
    print(f"{Colors.FAIL}No valid frames were processed from {carma_file}. Cannot calculate
averages.{Colors.ENDC}")
    return None

# Step 4: Package the results into the required data structure for the next function

# Create the mapping from the original formatted_atom_pairs index to the
# index in our unique_restraint_keys list. This is vital for the report generation.
column_mapping_original_to_filtered = {}
for original_idx, pair_desc in enumerate(formatted_atom_pairs):
    parts = pair_desc.split("****")
    detail_parts = parts[1].split("<==>")

```

```

left_psf = detail_parts[0].strip().split()
right_psf = detail_parts[1].strip().split(" (Degenerate)")[0].strip().split()

if len(left_psf) < 5 or len(right_psf) < 5: continue

atom1_key_tuple = (left_psf[2], left_psf[3], left_psf[4])
atom2_key_tuple = (right_psf[2], right_psf[3], right_psf[4])
current_restraint_key = tuple(sorted((atom1_key_tuple, atom2_key_tuple)))

if current_restraint_key in unique_restraint_keys:
    filtered_idx = unique_restraint_keys.index(current_restraint_key)
    column_mapping_original_to_filtered[original_idx] = filtered_idx

noe_data_results = {
    'nof_values': n_frames,
    'pairs_data': [],
    'filtered_keys_definition': unique_restraint_keys,
    'column_mapping_original_to_filtered': column_mapping_original_to_filtered
}

# Populate the pairs_data with the final sums
for i in range(num_unique_groups):
    pair_summary = {
        'filtered_col_index': i,
        'restraint_definition_key': unique_restraint_keys[i],
        'sum_r3': sum_r3[i],
        'sum_r6': sum_r6[i]
    }
    noe_data_results['pairs_data'].append(pair_summary)

```

```

return noe_data_results

def run_noe_averaging(noe_data, formatted_atom_pairs):
    """Generate NOE analysis report.
    noe_data: Contains sums (from process_noe_data) and mapping.
    formatted_atom_pairs: The original list from find_atom_ids_for_pairs, used to define unique display lines.
    """
    if not noe_data or not noe_data.get('pairs_data'):
        print(f"\n{Colors.FAIL}ERROR: No NOE data available to analyze.{Colors.ENDC}")
        return

    print(f"\n{Colors.HEADER}NOE Analysis Results:{Colors.ENDC}")
    print(f"{Colors.OKBLUE}{ '=' * 100}{Colors.ENDC}")
    print(f"{Colors.BOLD}  EXP      <r^-3>    <r^-6>    Deviation    Violation          Atoms\nInvolved{Colors.ENDC}")
    print(f"{Colors.OKBLUE}{ '=' * 100}{Colors.ENDC}")

    unique_display_entries = {}

    if not formatted_atom_pairs:
        print(f"{Colors.WARNING}formatted_atom_pairs is empty. Cannot generate detailed NOE analysis\n\n{Colors.ENDC}")
        return

    for original_idx, pair_desc_full_line in enumerate(formatted_atom_pairs):
        if "ERROR PARSING" in pair_desc_full_line:
            continue

```

```

try:
    initial_parts = pair_desc_full_line.split("****")
    if len(initial_parts) < 2:
        print(f"{Colors.WARNING}Malformed formatted_atom_pair line (missing '****'):"
              '{pair_desc_full_line}'. Skipping.{Colors.ENDC}")
        continue

    dist_part_split = initial_parts[0].strip().split()
    if len(dist_part_split) < 3:
        print(f"{Colors.WARNING}Malformed distance part in formatted_atom_pair line:"
              '{initial_parts[0].strip()}'. Skipping.{Colors.ENDC}")
        continue
    distance_str = dist_part_split[2]

    interaction_part_full = initial_parts[1].strip()
    is_line_marked_degenerate = "(Degenerate)" in interaction_part_full # Still track this for internal
    logic if needed later

    interaction_halves = interaction_part_full.split("<==>")
    if len(interaction_halves) < 2:
        print(f"{Colors.WARNING}Malformed interaction part (missing '<==>'): '{interaction_part_full}'.
              Skipping.{Colors.ENDC}")
        continue

    left_half_full_desc = interaction_halves[0].strip()
    right_half_full_desc_maybe_degen = interaction_halves[1].strip()
    right_half_full_desc = right_half_full_desc_maybe_degen.split(" (Degenerate)")[0].strip()

    l_parts = left_half_full_desc.split() # [ID, CHAIN, RESSEQ, RESNAME, ATOMNAME_MR]

```

```

r_parts = right_half_full_desc.split()

if len(l_parts) < 5 or len(r_parts) < 5:
    print(f"{Colors.WARNING}Could not parse atom details from: '{interaction_part_full}'. Expected 5
parts per side. Got L:{len(l_parts)}, R:{len(r_parts)}. Skipping.{Colors.ENDC}")
    continue

# Key components for display grouping: (ResSeq, ResName, AtomName_MR)
atom1_key_details = (l_parts[2], l_parts[3], l_parts[4])
atom2_key_details = (r_parts[2], r_parts[3], r_parts[4])
display_key = (distance_str, tuple(sorted((atom1_key_details, atom2_key_details))))

# Construct the shortened atom description for display
# Format: RESSEQ RESNAME ATOMNAME_MR
atom1_desc_short = f"{l_parts[2]:>4} {l_parts[3]:>3} {l_parts[4]:>4}"
atom2_desc_short = f"{r_parts[2]:>4} {r_parts[3]:>3} {r_parts[4]:>4}"
final_atom_details_for_print = f"{atom1_desc_short} <==> {atom2_desc_short}"

if display_key not in unique_display_entries:
    unique_display_entries[display_key] = {
        'atom_details_for_print': final_atom_details_for_print, # Store the shortened version
        'is_group_degenerate': is_line_marked_degenerate, # Store if *any* original line for this
group was degenerate
        'first_original_formatted_pair_idx': original_idx
    }
else:
    if is_line_marked_degenerate: # If any line contributing to this unique display is degenerate,
mark the group

```

```

        unique_display_entries[display_key]['is_group_degenerate'] = True

    except Exception as e:
        print(f"{Colors.FAIL}Error parsing formatted_atom_pair line for display: '{pair_desc_full_line}'.
Error: {e}{Colors.ENDC}")
        continue

stats = {
    'total_displayed_noes': 0,
    'violation_count': 0,
    'total_violation_sum': 0.0,
    'max_violation': 0.0,
    'by_severity': defaultdict(int)
}

for display_key, data in unique_display_entries.items():
    exp_distance = float(display_key[0])
    line_to_print_atom_part = data['atom_details_for_print']

    representative_original_idx = data['first_original_formatted_pair_idx']

    avg_r3_dist = 999.0
    avg_r6_dist = 999.0

    if noe_data.get('column_mapping_original_to_filtered') is None:
        print(f"{Colors.WARNING}column_mapping_original_to_filtered is missing in noe_data. Cannot map
restraints for averaging.{Colors.ENDC}")
    elif representative_original_idx in noe_data['column_mapping_original_to_filtered']:
        idx_in_noe_pairs_data = noe_data['column_mapping_original_to_filtered'][representative_original_idx]

```

```

if idx_in_noe_pairs_data < len(noe_data['pairs_data']):
    noe_sums = noe_data['pairs_data'][idx_in_noe_pairs_data]
    current_sum_r3 = noe_sums['sum_r3']
    current_sum_r6 = noe_sums['sum_r6']

    if noe_data['nof_values'] > 0:
        avg_r3_dist = (current_sum_r3 / noe_data['nof_values'])**(-1/3) if current_sum_r3 > 1e-9 else
999.0
        avg_r6_dist = (current_sum_r6 / noe_data['nof_values'])**(-1/6) if current_sum_r6 > 1e-9 else
999.0
    else:
        print(f"{Colors.WARNING}Mapped index {idx_in_noe_pairs_data} out of bounds for
noe_data.pairs_data (len {len(noe_data['pairs_data'])}). Restraint key for display: {display_key}{Colors.ENDC}")
    else:
        print(f"{Colors.WARNING}Original index {representative_original_idx} (key: {display_key}) not found
in noe_data's column_mapping. Cannot retrieve averaged distances.{Colors.ENDC}")

deviation = avg_r6_dist - exp_distance if avg_r6_dist != 999.0 else 999.0
violation = max(0, deviation) if deviation != 999.0 else 999.0

stats['total_displayed_noes'] += 1
if violation > 1e-3 and violation != 999.0:
    stats['violation_count'] += 1
    stats['total_violation_sum'] += violation
    stats['max_violation'] = max(stats['max_violation'], violation)

if violation > 3.0: stats['by_severity']['>3.0Å'] += 1
elif violation > 2.0: stats['by_severity']['>2.0Å'] += 1

```

```

        elif violation > 1.0: stats['by_severity']['>1.0Å'] += 1
        elif violation > 0.5: stats['by_severity']['>0.5Å'] += 1
        else: stats['by_severity']['0-0.5Å'] +=1

v_color = Colors.ENDC
if violation == 999.0: v_color = Colors.FAIL
elif violation > 3.0: v_color = Colors.FAIL
elif violation > 1.0: v_color = Colors.WARNING
elif violation > 1e-3 : v_color = Colors.OKGREEN

print(
    f" {exp_distance:5.2f}    {avg_r3_dist:6.3f}    {avg_r6_dist:6.3f}    "
    f"{deviation:7.3f}    {v_color}{violation:7.3f}{Colors.ENDC}    "
    f"{line_to_print_atom_part}"
)

print(f"\n{Colors.OKBLUE}{'=' * 100}{Colors.ENDC}")
if stats['total_displayed_noes'] > 0:
    violation_pct = (stats['violation_count'] / stats['total_displayed_noes']) * 100 if
stats['total_displayed_noes'] > 0 else 0
    avg_viol = stats['total_violation_sum'] / stats['violation_count'] if stats['violation_count'] > 0 else
0.0

summary_color = Colors.OKGREEN
if violation_pct > 30: summary_color = Colors.FAIL
elif violation_pct > 15: summary_color = Colors.WARNING

print(f"{Colors.BOLD}SUMMARY: {stats['total_displayed_noes']} unique NOE restraints
displayed.{Colors.ENDC}")

```

```

        print(f"{summary_color}Violations (>0.001Å): {stats['violation_count']}
({violation_pct:.1f}%) {Colors.ENDC}")
        print(f"Average violation (for violated restraints): {avg_viol:.3f} Å")
        print(f"Maximum violation: {stats['max_violation']:.3f} Å")
        print("Violation severity distribution:")
        severity_order = ['0-0.5Å', '>0.5Å', '>1.0Å', '>2.0Å', '>3.0Å']
        sorted_severity_keys = [key for key in severity_order if key in stats['by_severity']]
        for severity_range in sorted_severity_keys:
            print(f"  {severity_range}: {stats['by_severity'][severity_range]} restraints")
        for sev_key, count in stats['by_severity'].items():
            if sev_key not in sorted_severity_keys:
                print(f"  {sev_key}: {count} restraints")
    else:
        print(f"{Colors.WARNING}No NOEs were displayed in the analysis.{Colors.ENDC}")

def identify_files(args):
    """Identify input files by extension"""
    mr_file = psf_file = dcd_file = None
    for arg in args:
        if arg.lower().endswith('.mr'):
            mr_file = arg
        elif arg.lower().endswith('.psf'):
            psf_file = arg
        elif arg.lower().endswith('.dcd'):
            dcd_file = arg
    if not mr_file or not psf_file:
        print(f"{Colors.FAIL}Error: Both .mr and .psf files are required.{Colors.ENDC}")
        sys.exit(1)

```

```

    return mr_file, psf_file, dcd_file

def main():
    if len(sys.argv) < 3:
        print(f"{Colors.FAIL}Usage: python3 {sys.argv[0]} <mr_file> <psf_file> [<dcd_file>]{Colors.ENDC}")
        sys.exit(1)

    mr_file_path, psf_file_path, dcd_file_path = identify_files(sys.argv[1:])

    print(f"{Colors.OKCYAN}Processing MR file: {mr_file_path}, PSF file: {psf_file_path}{Colors.ENDC}")
    if dcd_file_path: print(f"{Colors.OKCYAN}DCD file provided: {dcd_file_path}{Colors.ENDC}")

    # Generate the atom pair list, which includes experimental distances and detailed atom info.
    # This list is used throughout the script to define restraints.
    formatted_atom_pairs, has_unknown = find_atom_ids_for_pairs(mr_file_path, psf_file_path)

    print(f"\n{Colors.OKBLUE}Atom Pair Table (from MR and PSF):{'='*43}{Colors.ENDC}")
    if not formatted_atom_pairs:
        print(f"{Colors.WARNING}No atom pairs could be formatted. Check MR/PSF files and parsing logic.{Colors.ENDC}")

    # Print the table to the console (without experimental distance)
    for line in formatted_atom_pairs:
        line_to_print_console = line
        if "****" in line:
            parts = line.split("****", 1)
            first_part_elements = parts[0].strip().split()
            if len(first_part_elements) == 3:
                line_to_print_console = f"{first_part_elements[0]:>4}"

```

```

{first_part_elements[1]:>4}{'':4s}***{parts[1]}"

    if "???" in line or "ERROR" in line.upper():
        print(f"{Colors.FAIL}{line_to_print_console}{Colors.ENDC}")
    else:
        print(line_to_print_console)
print(f"{Colors.OKBLUE}{'='*77}{Colors.ENDC}")

table1_out_path = "table1.out"
with open(table1_out_path, "w") as table_file:
    table_file.write("="*85 + "\n")
    for line in formatted_atom_pairs:
        line_to_write_file = line
        if "****" in line:
            parts = line.split("****", 1)
            first_part_elements = parts[0].strip().split()
            if len(first_part_elements) == 3:
                line_to_write_file = f"{first_part_elements[0]:>4}
{first_part_elements[1]:>4}{'':4s}***{parts[1]}"
                table_file.write(line_to_write_file + "\n")
            table_file.write("="*85 + "\n")
    print(f"\n{Colors.OKGREEN}Full atom pair table saved to {table1_out_path}{Colors.ENDC}")

if has_unknown:
    print(f"\n{Colors.FAIL}Unknown atoms or parsing errors encountered during PSF lookup. Aborting before
CARMA.{Colors.ENDC}")
    sys.exit(1)
if not formatted_atom_pairs:
    print(f"\n{Colors.FAIL}No valid atom pairs generated. Aborting before CARMA.{Colors.ENDC}")

```

```

sys.exit(1)

carma_input_dist_file = "carmainput_for_distcalc.txt"
try:
    # Write only the unique atom ID pairs for CARMA's -dist flag.
    with open(carma_input_dist_file, "w") as f_carma_in:
        unique_id_pairs_for_carma = set()
        for line in formatted_atom_pairs:
            if "ERROR" in line.upper() or "???" in line: continue
            parts = line.split()
            if len(parts) >= 2:
                try:
                    id_pair = tuple(sorted((int(parts[0]), int(parts[1]))))
                    if id_pair not in unique_id_pairs_for_carma:
                        f_carma_in.write(f"{parts[0]} {parts[1]}\n")
                        unique_id_pairs_for_carma.add(id_pair)
                except ValueError:
                    pass # Silently skip malformed lines

carma_distances_file = "carma.distances"

if dcd_file_path:
    run_carma_calc = True
    if os.path.exists(carma_distances_file):
        resp = input(f"\n{Colors.WARNING}File '{carma_distances_file}' already exists. Regenerate with
CARMA? (y/n, default y): {Colors.ENDC}").lower().strip()
        if resp in ('n', 'no'):
            run_carma_calc = False
            print(f"{Colors.OKCYAN}Skipping CARMA calculation, using existing

```

```

'{carma_distances_file}'.{Colors.ENDC}")

    if run_carma_calc:
        print(f"\n{Colors.OKBLUE}Starting CARMA calculation at
{datetime.datetime.now().strftime('%Y-%m-%d %H:%M:%S')}{Colors.ENDC}")
        # Use the generated input file for the -dist argument
        carma_cmd = ["carma", "-v", "-atmid", "ALLID", "-dist", carma_input_dist_file, psf_file_path,
dcd_file_path]
        print(f"Executing: {' '.join(carma_cmd)}")
        try:
            subprocess.run(carma_cmd, check=True)
            print(f"{Colors.OKGREEN}CARMA calculation finished successfully at
{datetime.datetime.now().strftime('%Y-%m-%d %H:%M:%S')}{Colors.ENDC}")
        except subprocess.CalledProcessError as e:
            print(f"{Colors.FAIL}CARMA execution failed with error: {e}{Colors.ENDC}")
            sys.exit(1)
        except FileNotFoundError:
            print(f"{Colors.FAIL}CARMA command not found. Please ensure CARMA is installed and in your
PATH.{Colors.ENDC}")
            sys.exit(1)

# Proceed with analysis if the distance file exists.
if os.path.exists(carma_distances_file):
    print(f"\n{Colors.OKBLUE}Starting NOE processing using '{carma_distances_file}' at
{datetime.datetime.now().strftime('%Y-%m-%d %H:%M:%S')}{Colors.ENDC}")

# The new function combines filtering and processing. No user input is needed.
# It takes the full atom pair list to understand the structure of the raw distance file.
noe_data_struct = calculate_noe_sums_in_memory(formatted_atom_pairs, carma_distances_file)

```

```

    if noe_data_struct:
        # The final report generation uses the processed sums.
        run_noe_averaging(noe_data_struct, formatted_atom_pairs)
    else:
        print(f"{Colors.FAIL}NOE data processing failed. Cannot run averaging.{Colors.ENDC}")

        print(f"\n{Colors.OKGREEN}NOE processing completed at {datetime.datetime.now().strftime('%Y-%m-%d
%H:%M:%S')}{Colors.ENDC}")

    elif dcd_file_path:
        print(f"{Colors.FAIL}'{carma_distances_file}' is missing. Cannot proceed.{Colors.ENDC}")
    else:
        print(f"{Colors.WARNING}No DCD file provided and no existing '{carma_distances_file}'. Cannot
perform NOE analysis.{Colors.ENDC}")

    finally:
        # Clean up the temporary input file for CARMA.
        if os.path.exists(carma_input_dist_file):
            try:
                os.remove(carma_input_dist_file)
            except OSError as e:
                print(f"{Colors.WARNING}Could not remove temporary file {carma_input_dist_file}:
{e}{Colors.ENDC}")

if __name__ == "__main__":
    main()

```


References

- Brooks, B. R., Brooks, C. L., III, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., ... & Karplus, M. (2009). CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*, 30(10), 1545–1614. <https://doi.org/10.1002/jcc.21287> [1]
- Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., & Shaw, D. E. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics*, 78(8), 1950–1958. <https://doi.org/10.1002/prot.22711> [1]
- Honda, S., Akiba, T., Kato, Y. S., Sawada, Y., Sekijima, M., Ishimura, M., Ooishi, A., Watanabe, H., Odahara, T., & Harata, K. (2008). Crystal structure of a ten-amino acid protein. *Journal of the American Chemical Society*, 130(46), 15327–15331. <https://doi.org/10.1021/ja8030533>
- Honda, S., Yamasaki, K., Sawada, Y., & Morii, H. (2004). 10 residue folded peptide designed by segment statistics. *Structure (London, England : 1993)*, 12(8), 1507–1518. <https://doi.org/10.1016/j.str.2004.05.022>
- Glykos, N.M. (2006), Software news and updates carma: A molecular dynamics analysis program. *J. Comput. Chem.*, 27: 1765-1768. <https://doi.org/10.1002/jcc.20482>
- Koukos, P. I., & Glykos, N. M. (2013). Grcarma: A fully automated task-oriented interface for the analysis of molecular dynamics trajectories. *Journal of computational chemistry*, 34(26), 2310–2312. <https://doi.org/10.1002/jcc.23381>
- Athanasia-Panagiota Serafeim, Georgios Salamanos, Kalliopi K. Patapati, and Nicholas M. Glykos (2016), *Journal of Chemical Information and Modeling* 2016 56 (10), 2035-2041. <https://doi.org/10.1021/acs.jcim.6b00493>
- Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., & Wüthrich, K. (1998). Recommendations for the presentation of NMR structures of proteins and nucleic acids. IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy. *Journal of biomolecular NMR*, 12(1), 1–23. <https://doi.org/10.1023/a:1008290618449>
- Zirafi, O., Kim, K. A., Ständker, L., Mohr, K. B., Sauter, D., Heigele, A., ... & Münch, J. (2015). Discovery and characterization of an endogenous CXCR4 antagonist. *Cell Reports*, 11(5), 737–747. <https://doi.org/10.1016/j.celrep.2015.03.061>
- Neuhaus, D. (2011). Nuclear Overhauser Effect. In *eMagRes* (eds R.K. Harris and R.L. Wasylishen). <https://doi.org/10.1002/9780470034590.emrstm0350.pub2>
- Williamson, M.P. (2008). The Nuclear Overhauser Effect. In: Webb, G.A. (eds) *Modern Magnetic Resonance*. Springer, Dordrecht. https://doi.org/10.1007/1-4020-3910-7_50
- V. Sánchez, A.G. Redfield, P.D. Johnston, & J. Tropp (1980), Nuclear Overhauser effect in specifically deuterated macromolecules: NMR assay for unusual base pairing

- in transfer RNA., *Proc. Natl. Acad. Sci. U.S.A.* 77 (10) 5659-5662, <https://doi.org/10.1073/pnas.77.10.5659> .
- Wolfgang Rieping, Michael Habeck, Benjamin Bardiaux, Aymeric Bernard, Thérèse E. Malliavin, Michael Nilges, ARIA2: Automated NOE assignment and data integration in NMR structure calculation, *Bioinformatics*, Volume 23, Issue 3, February 2007, Pages 381–382, <https://doi.org/10.1093/bioinformatics/btl589>
 - Herrmann, T., Güntert, P., & Wüthrich, K. (2002). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *Journal of molecular biology*, 319(1), 209–227. [https://doi.org/10.1016/s0022-2836\(02\)00241-3](https://doi.org/10.1016/s0022-2836(02)00241-3)
 - Montelione, G., Zheng, D., Huang, Y. *et al.* Protein NMR spectroscopy in structural genomics. *Nat Struct Mol Biol* 7 (Suppl 11), 982–985 (2000). <https://doi.org/10.1038/80768>
 - Harris, C.R., Millman, K.J., van der Walt, S.J. et al. Array programming with NumPy. *Nature* 585, 357–362 (2020). <https://doi.org/10.1038/s41586-020-2649-2>
 - Daura, X., Antes, I., van Gunsteren, W.F., Thiel, W. and Mark, A.E. (1999), The effect of motional averaging on the calculation of NMR-derived structural properties. *Proteins*, 36: 542-555. [https://doi.org/10.1002/\(SICI\)1097-0134\(19990901\)36:4](https://doi.org/10.1002/(SICI)1097-0134(19990901)36:4)
 - Peter, C., Daura, X., & van Gunsteren, W. F. (2001). Calculation of NMR-relaxation parameters for flexible molecules from molecular dynamics simulations. *Journal of biomolecular NMR*, 20(4), 297–310. <https://doi.org/10.1023/a:1011241030461>
 - Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*. John Wiley & Sons.
 - Günther, H. (2013). *NMR spectroscopy: basic principles, concepts and applications in chemistry*. John Wiley & Sons.
 - Keeler, J. (2010). *Understanding NMR spectroscopy*. John Wiley & Sons.