Data Article

Development and benchmark to obtain AMBER parameters dataset for non-standard amino acids modified with 4-hydroxy-2-nonenal

Antistio Alviz-Amador a, Rodrigo Galindo-Murillo b, Rafael Pineda-Alemán c, Humberto Pérez-González d, Erika Rodríguez-Cavallo a, Ricardo Vivas-Reyes e, Darío Méndez-Cuadro c,*

a Analytical Chemistry and Biomedicine Group, Pharmaceutical Sciences Faculty, University of Cartagena, Cartagena, Colombia
b Department of Medicinal Chemistry, University of Utah, Salt Lake, USA
c Analytical Chemistry and Biomedicine Group, Exact and Natural Sciences Faculty, University of Cartagena, Cartagena, Colombia
d Department of Mathematical, Exact and Natural Sciences Faculty, University of Cartagena, Cartagena, Colombia
e Grupo de Química Cuántica y Teórica, Facultad de Ciencias Exactas y Naturales, Universidad de Cartagena, Cartagena, Colombia

A R T I C L E   I N F O

Article history:
Received 2 November 2018
Received in revised form 16 November 2018
Accepted 20 November 2018
Available online 27 November 2018

Keywords:
AMBER
Gaff
Force field parameterization
Mechanical quantic
Molecular dynamics
Geometry optimization
Validation

A B S T R A C T

The data described here support the research article “4-HNE carbonylation induces local conformational changes on bovine serum albumin and thioredoxin. A molecular dynamics study” (Alviz-Amador et al., 2018). Dataset on Gaff force field parameters of AMBER is provided for assembled three non-standard amino acids resulting of the 4-HNE Michael addition, the main end product of lipids peroxidation. Data include a framework for derivation of missing bonds, angles and dihedral parameters for Cys, His, and Lys modified amino acids, alongside optimized partial charges derived with Restrained Electrostatic Potential (RESP) method and the new force field parameters obtained by quantic mechanical (QM) using HF/6-31G** level of theory. Benchmark as a graphics tutorial summary steps to obtained new parameters and the validation of non-standard amino acids is presented. The new residues...
Specifications table

| Subject area                  | Biochemistry, Biophysics |
|-------------------------------|--------------------------|
| More specific subject area    | Computational Biochemistry, Computational Biophysics |
| Type of data                  | Figures and tables |
| How data were acquired        | Quantum Mechanics (QM), Molecular Dynamics (MD), Software used: Gaussian 09 for QM, AMBER (pmmemd) for MD |
| Data format                   | Raw and analyzed. |
| Experimental factors          | (2S, 4S, 5R)-4-Hydroxy-2-nonenal isomer was used to build Michael adducts |
| Experimental features         | Theoritical level HF/6-31G** for QM and Gaff2 force field and ff14SB force field for MD |
| Data source location          | Cartagena, Colombia, Facultad de Ciencias Farmacéuticas and Facultad de Ciencias Exactas y Naturales.10°23'58. 75°30'09., Cl. 6 #3°N, Cartagena, Bolivar |
| Data accessibility            | Data are supplied with this article. Parameter files are available http://research.bmh.manchester.ac.uk/bryce/amber/ |
| Related research article      | A. Alviz-Amador, R. Galindo-Murillo, R. Pineda-Alemán, H. Pérez-González, E. Rodríguez-Cavallo, R. Vivas-Reyes, D. Méndez-Cuadro. 4-HNE carbonylation induces local conformational changes on bovine serum albumin and thioredoxin. A molecular dynamics study, J. Mol. Graph. Model. (2018). doi:10.1016/J.JMGM.2018.11.001. [1] |

Value of the data

- Dataset of new AMBER force field parameters are provided to perform Molecular Dynamics Simulation of 4-HNE carbonylated proteins with Michael adducts on Cys, His and Lys residues.
- A benchmark framework for constructing, parameterizing, optimizing and validating of new non-standard amino acid residues modified with 4-HNE is now available.
- Our data can be used to modify, simulate and evaluate by molecular dynamic simulation the effects of 4-HNE carbonylation over any protein system.

1. Data

In the Supplementary material the dataset of partial charges assigned to Cys-HNE, His-HNE and Lys-HNE are shown. In Fig. 1, the workflow for preparing parameter files for non-standard residues is described. In addition, in Fig. 2 the optimized structures obtained with theory level HF/6-31G** are presented. Finally, in Tables 1–3 the dataset with the information of new obtained parameters are listed as coordinates files for Cys-HNE, His-HNE and Lys-HNE, respectively.
2. Experimental design, materials, and methods

2.1. Parameterization

Dataset of Gaff force field parameters were established for three non-standard amino acids (His-HNE, Lys-HNE and Cys-HNE) to be used for molecular dynamics simulations of proteins [1] The framework for derivation of missing bond, angle and dihedral parameters is presented in Fig. 1. First, non-standard amino acid were constructed with GaussView 5, followed by full geometry optimization of the new structures using the Hartree–Fock level (HF/6-31G**) [2]. Next, assignment of charges, missing bonds, angles, and dihedral angles parameters were constructed with the antechamber and leap programs as included in AmberTools [3]. Then, charges (Step 4) of the optimized structures were calculated using RESP method [4] and the partial charges assigned to individual atoms are listed in the Supplementary material. Missing bonds, angles, and dihedral parameters of each amino acids modified with 4-HNE were established by homology, matching atom types automatically from the Gaff force field and using parmchk to generate the required force constants [5]. Dataset of new parameters assigned for the new non-standard amino acids were consigned in frcmod files and they are summarized below Tables 1–3. Next, coordinate and topology files were created for each non-standard amino acid with the program leap.

These non-standard amino acids were replaced on the proteins and the lacking parameters in frcmod files corresponding to peptide bonds, angle and torsions between the non-standard amino acids and the end nitro-terminus and the end carboxyl terminus of the nearby amino acids on proteins, were calculated using the program parmcal of Antechamber package. These improved frcmod files were loaded into leap program from AmberTools16 to generate the libraries files (type lib files).

Finally, the optimized structures of modified and unmodified amino acids Cys, His and Lys with Michael adducts are showed in Fig. 2; whereas the new improved parameters were included into Tables 1–3. There, bond parameters values are expressed as bond constants \( (k_r) \) in kcal mol\(^{-1}\) Å\(^{-2}\); distance at equilibrium \( (r_{eq}) \) in Å; angle constant \( (k_\theta) \) in kcal mol\(^{-1}\) degree\(^{-2}\); angle at equilibrium \( (\theta_{eq}) \) in degrees, dihedrals constants \( (V_{n/2}) \) in kcal/mol and dihedrals constants angles \( (\psi) \) in degrees.
Fig. 2. Optimized structures of non-modified and modified amino acids Cys, His and Lys with Michael adducts of 4-HNE. Figures were obtained with theory level HF/6-31G** and the atom names follows PDB conventions.
### Table 1
New parameters assigned to CYS-4HNE.

| Bond | Atomtypes | Kr | req |
|------|-----------|----|-----|
| c1-N | 361.821   | 1.431 |

| Angle | Atomtypes | ΘKr | Θeq | Note |
|------|-----------|-----|-----|------|
| n2-c3-c1 | 68.008 | 109.895 | Calculated with empirical approach |
| c3-c1-o | 57.930 | 119.420 | Calculated with empirical approach |
| c1-N-H | 50.000 | 120.000 |
| c1-N-CX | 63.311 | 133.250 |
| c3-c1-N | 66.420 | 114.600 |

| Dihedral | Atomtypes | Vn/2 | γ (ψ) | n |
|----------|-----------|------|-------|---|
| o-c1-N-H | 1 | 180 | 2 |
| o-c1-N-CX | 1 | 180 | 2 |
| c3-c1-N-H | 1 | 180 | 2 |
| c3-c1-N-CX | 1 | 180 | 2 |

### Table 2
New parameters assigned to HIS-4HNE.

| Bond | Atomtypes | Kr | req |
|------|-----------|----|-----|
| C-n2 | 361.821 | 1.431 |
| c1-N | 361.821 | 1.431 |

| Angle | Atomtypes | ΘKr | Θeq | Note |
|------|-----------|-----|-----|------|
| O-C-n2 | 72.926 | 119.420 |
| C-n2-c3 | 63.311 | 133.250 |
| C-n2-hn | 46.635 | 118.320 |
| CX-C-n2 | 66.420 | 114.600 |
| o-c1-N | 72.926 | 119.420 |
| c1-N-CT | 50.000 | 121.900 |
| c1-N-CX | 63.311 | 133.250 |
| c3-c1-N | 66.420 | 114.600 |
| n2-c3-c1 | 68.008 | 109.895 | Calculated with empirical approach |
| c3-c1-o | 57.930 | 119.420 | Calculated with empirical approach |

| Dihedral | Atomtypes | Vn/2 | γ (ψ) | n |
|----------|-----------|------|-------|---|
| O-C-n2-c3 | 1 | 180 | 2 |
| O-C-n2-hn | 1 | 180 | 2 |
| CX-C-n2-c3 | 1 | 180 | 2 |
| CX-C-n2-hn | 1 | 180 | 2 |
| o-c1-N-CT | 1 | 180 | 2 |
| o-c1-N-CX | 1 | 180 | 2 |
| c3-c1-N-CT | 1 | 180 | 2 |
| c3-c1-N-CX | 1 | 180 | 2 |

| Improper | Atomtypes | Vn/2 | γ (ψ) | n | Note |
|----------|-----------|------|-------|---|------|
| h5-n2-c2-na | 1,1 | 180.0 | 2.0 | Using default value |
| c2-h4-c2-na | 1,1 | 180.0 | 2.0 | Using default value |
From these datasets, the topology and coordinate of modified proteins were obtained. Hence, the applicability of the newly derived MM parameter, they were subsequently employed in 1 µs MD simulations of each of the three non-natural amino acids treated following the methodology described by [1] and [6].

2.2. Validation

To test the generated structures from the modified amino acids we performed MD simulations as described above using only the modified structure and compared selected bond distances and angles with structures obtained from DFT level of theory m062x/631g (d) (Table 4). Overall, good agreement between the data from high-level QM calculations and the generated AMBER structures were seen. Distance average error is between the ranges of 0.02–0.05 Å whereas angle error is within 4°. As an additional comparison, we extracted the bonds and angles information from our modified protein systems and found that average bond difference is within 0.01–0.03 Å and within 3.7°.

2.3. Analysis of molecular dynamics trajectories of non-standard vs. standard amino acids

All atom root means square deviation analysis for unmodified and modified amino acids is presented in Fig. 3. Distance found in RMSD analysis for all unmodified amino acids were under 1 Å, being 0.8 Å for Hys and Lys and 0.4 Å on Cys. In the case of non-standard amino acids, RMSD values increased until an average of 1.4 Å for His-4HNE and Lys-4HNE and average of 1.2 Å for Cys-HNE (Fig. 3). Differences observed fall into a range of 0.6–0.8 Å for RSMD comparisons among modified/unmodified amino acids indicating that 4-HNE do not induce dramatically structural changes.
Table 4
Comparison between selected bond distances and angles calculated from optimized nonstandard amino acids structures. Data from the single modified amino acids were extracted from a 1 µs MD simulation using the same protocols describe before, comparisons were calculated using the DFT level of theory m062x and a basis set 6–31g (d).

| Validation methods | CYS-4HNE | HIS-4HNE | LYS-4HNE |
|--------------------|---------|----------|----------|
| QM (m062x/631g (d)) | Bond (Å, ± Stdev) | Angle (°, ± Stdev) | Bond (Å, ± Stdev) | Angle (°, ± Stdev) | Bond (Å, ± Stdev) | Angle (°, ± Stdev) |
| SG-C | 1.82 | ± 0.05 | NE2-C | 1.45 | ± 0.03 | NZ-C | 1.45 | ± 0.03 |
| C4-C5 | 1.52 | ± 0.03 | C4-C5 | 1.53 | ± 0.01 | C4-C5 | 1.52 | ± 0.03 |
| SG-C-C1 | 115.25 | ± 4.01 | NE2-C-C1 | 112.07 | ± 5.17 | C4-C5-C6 | 112.73 | ± 3.3 |
| C4-C5-C6 | 112.37 | ± 4.06 | C4-C5-C6 | 112.8 | ± 0.97 | NE2-C-C1 | 112.07 | ± 3.3 |
| MM (AMBER) aa alone | 1.83 | ± 0.01 | 1.47 | ± 0.01 | 1.50 | ± 0.03 | 1.48 | ± 0.03 |
| 1.56 | ± 5.17 | 1.54 | ± 5.17 | 1.55 | ± 3.3 | 1.54 | ± 5.17 |
| 111.77 | ± 4.06 | 112.22 | ± 0.97 | 117.14 | ± 4.06 | 113.40 | ± 0.97 |
| 112.56 | ± 4.06 | 113.23 | 112.37 | 113.40 |
| MM (AMBER) aa on protein | 1.50 | ± 0.03 | 1.48 | ± 0.03 | 1.50 | ± 0.03 | 1.48 | ± 0.03 |
| 1.54 | ± 3.3 | 1.55 | ± 3.3 | 1.56 | ± 3.3 |
| 113.41 | ± 0.97 | 117.14 | ± 0.97 | 112.73 | ± 0.97 |
| 113.40 | ± 0.97 | 113.40 | 113.40 |

Atoms involved in selected distances and angles.
Acknowledgments

We thank the computer time provided by the Center for High Performance Computing at the University of Utah and Professor Thomas E. Cheatham III for helpful discussion and support. This work was supported by research grant code 1107-712-50088 from Administrative Department of Science, Technology and Research - COLCIENCIAS and the University of Cartagena, (Colombia). AAA was supported by COLCIENCIAS, Colfuturo and the University of Cartagena, Doctoral scholarship 757-2016. DMC, ERC and RVR are indebted to the University of Cartagena, for the continuous support to research groups.

Transparency document. Supporting information

Transparency document associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2018.11.102.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2018.11.102.

References

[1] A. Alviz-Amador, R. Galindo-Murillo, R. Pineda-Alemán, H. Pérez-González, E. Rodríguez-Cavallo, R. Vivas-Reyes, D. Méndez-Cuadro, 4-HNE carbonylation induces local conformational changes on bovine serum albumin and thioredoxin. A molecular dynamics study, J. Mol. Graph. Model. 86 (2019) 298–307, https://doi.org/10.1016/J.JMGM.2018.11.001.
[2] W.J. Hehre, Ab initio molecular orbital theory, Acc. Chem. Res. 9 (1976) 399–406. https://doi.org/10.1021/ar50107a003.
[3] D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, Amber 2016 Reference Manual, 2016. (https://doi.org/10.1021/ct200909j).

[4] W.D. Cornell, P. Cieplak, C.I. Bayly, P.A. Kollman, Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation, J. Am. Chem. Soc. 115 (1993) 9620–9631. https://doi.org/10.1021/ja00074a030.

[5] E.M. Gardner, M.E. Maravi, C. Rietmeijer, J. Arthur, W.J. Burman, D.P. Health, The amber biomolecular simulation programs, J. Comput. Chem. 6 (2009) 145–155. https://doi.org/10.1038/nm.2120.Red-shifted.

[6] O.O. Olubiyi, B. Strod, Topology and parameter data of thirteen non-natural amino acids for molecular simulations with CHARMM22, Data Brief 9 (2016) 642–647. https://doi.org/10.1016/j.dib.2016.09.051.