Lawrence Berkeley National Laboratory

LBL Publications

Title

Accurate geometries for “Mountain pass” regions of the Ramachandran plot using quantum chemical calculations

Permalink

https://escholarship.org/uc/item/9mj8d628

Journal

Proteins Structure Function and Bioinformatics, 86(3)

ISSN

0887-3585

Authors

Jiang, Zhongming
Biczysko, Malgorzata
Moriarty, Nigel W

Publication Date

2018-03-01

DOI

10.1002/prot.25451

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Accurate geometries for “mountain pass” regions of the Ramachandran plot using quantum chemical calculations

Short title: Ramachandran plot “mountain pass” regions

Keywords: protein, poly-peptides, unusual arrangements, DFT, B3LYP-D3, backbone, local structure

Zhongming Jiang[1], Malgorzata Biczysko[1], Nigel W. Moriarty[2]

[1] International Centre for Quantum and Molecular Structures, College of Sciences, Shanghai University, 99 Shangda Road, Shanghai 200444, China

[2] Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

The institution at which the work was performed:
International Centre for Quantum and Molecular Structures, College of Sciences, Shanghai University, 99 Shangda Road, Shanghai 200444, China

Corresponding Author:

Nigel W. Moriarty

Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Correspondence email: nwmoriarty@lbl.gov

Abstract

Unusual local arrangements of protein in Ramachandran space is not well represented by standard geometry tools used in either protein structure refinement using simple harmonic geometry restraints or in protein simulations using molecular mechanics force fields. By contrast, quantum chemical computations using small poly-peptide molecular models can predict accurate geometries for any well-defined backbone Ramachandran orientation. For conformations along transition regions – ϕ from -60 to 60° – a very good agreement with representative high-resolution experimental X-ray (≤1.5 Å) protein structures is obtained for both backbone C^{-1}-N-Cα angle and the nonbonded O^{-1}…C distance, while "standard geometry" leads to the "clashing" of O…C atoms and Amber FF99SB predicts distances too large by about 0.15 Å. These results confirm that quantum chemistry computations add valuable support for detailed analysis of local structural arrangements in proteins, providing improved or missing data for less understood high-energy or unusual regions.

Introduction

Extensive analysis of the peptide backbone dihedral angles in the structures taken from the Protein Data Bank (PDB) [1] is concentrated on the energetically favored minima in the Ramachandran (ϕ, ψ) plots [2]. This is because these are the most common arrangements for which increased number of examples improves the accuracy of the statistical sampling of structural parameters. This reference data is used to guide and validate structure building and refinement in the X-ray crystallography field. It is common for proteins with peptides that have Ramachandran outliers to be scrutinized in order to bring the outliers into a favored or allowed region of the plot [3,4]. Of course, outliers are not necessarily wrong but are more likely to be so. This outlier status makes it less likely that the more detailed analysis of the bond lengths and bond angles has been performed. This can hinder reproduction of the geometry with either a force field or a simple restraints approach commonly used in X-ray refinement of proteins, which both rely on prior knowledge. This is because both have been developed using the curated data that confirms to the vast majority of protein configurations in the hope that it will perform in other situations as well and guide the protein to a global minimum.

It should be noted that one of the features of X-ray crystallography protein refinement is that the restraints are weighted against the experimental data to provide the best possible model [5]. In the high-resolution cases considered in this analysis, the geometry restraints are down weighted allowing the experimental information to drive the refinement. This means that high-resolution structures are less reliant on the prior of the restraints and can be mined for better restraints. By contrast, in low-resolution cases the restraints play a larger role to maintain a reasonable geometry, due to the lack of sufficiently detailed experimental data. If the geometry model does

not allow for atypical configurations, the refinement will be driven into an allowed conformation that is not correct.

The geometry data gleaned from the high-resolution structures is rendered in the restraints model [6,7]. Each of the internal coordinates of a molecule has a corresponding restraint with the potential represented by a quadratic curve. The minimum of the curve is the ideal value of the internal coordinate and the estimated standard deviation of restraint is used to determine the width of the curve and, in turn, the gradients towards the ideal value. Each particular internal coordinate in a protein model is set to a certain value defining the "standard geometry". For instance the C_α–C_β bond in an alanine is 1.521Å, regardless of the configuration of the peptide.

The restraints approach can be improved using a dynamic lookup of the bond lengths and bond angles to use for restraints in refinement based on the (ϕ, ψ) values of the peptide. The conformation dependent library (CDL) [8-10] greatly improves the agreement of the geometry restraints to the final protein geometry in the majority of cases. The CDL has been constructed based on data gleaned from high-resolution structures. The reference values included in CDL show significant deviations, with some backbone angles varying up to 6 degrees even for (ϕ, ψ) in the allowed regions. Yet, still constant values of restraints are associated to the large portion of the Ramachandran space due to the lack of reliable conformation-dependent reference parameters. It is not hard to see why standard restraints fail (even including CDL) in genuine cases of atypical configurations that are outside the allowed regions but are nonetheless legitimate.

One such example of atypical configurations was found in a recent analysis of 616,212 non-glycine residues from representative, better than 1.5 Å resolution structures. It revealed 146 high-energy, yet stable, conformations with the Ramachandran angles in the range $-35° < \varphi < 35°$, and

the ψ about 90 and -90°. These φ ~ 0 mountain pass residues, previously expected to occur only as transition states, are stably trapped in native protein structures [11]. An example, the Alanine (Ala:12) residue from the PDB entry 3VUR is shown in Figure 1a, while all 146 residues are depicted as blue dots on Ramachandran plot in Figure 2a. Analysis of the PDB structures gathered detailed experimental information on geometries from conformational transition regions. This database of the unusual backbone conformations proved to be a hard test for tools commonly applied in protein science [11]. Indeed unusual local arrangements are not well described, not only by the simple geometry restraints used in protein structure refinement, but also by the state-of the-art molecular mechanics force-fields, such as AMBER FF99SB [12,13], which was considered as giving the best performance in a protein modeling test [14]. The latter, may have important implications for the analysis of protein active sites as well as for the validity of molecular dynamics studies related to the conformational switching between main Ramachandran minima basins [11]. For instance relaxation timescales and the conformational exchange processes derived from simulations with each of the four major force field families (including the AMBER FF99SB) differ significantly [15]. We demonstrate how these problems can be overcome by using quantum mechanical (QM) computations, with methodologies that are proven to deliver accurate structures and properties of isolated polypeptides [16-18]. Similar approaches using first-principle based data sets have been reported for isolated and cation-coordinated conformers of proteinogenic amino acids [19]. In this work the focus is on structural parameters within the O^{-1}-C^{-1}-N-Cα-C^{1} ring of high-energy “mountain regions” described above (see Figure 2). However, considering advances in quantum chemistry for the description of all types of weak molecular interactions [20,21] and related three-dimensional structures [21,22], similar accuracy can be expected for other non-typical local arrangements [23]. There are several

possible applications of quantum-chemistry based approaches, either directly or as reference data for less-expensive methodologies. Most common and explored examples are represented by detailed studies of interesting regions or active sites [24], or the molecular mechanics force fields optimization providing improved parameters [25,26]. Here we highlight other already feasible yet less-regarded applications to the protein crystallographic refinement, in particular a definition of conformation-dependent restraints [8,10]. However, direct combination of computational and experimental data in quantum refinement procedures is also becoming possible [27,28].

Materials and Methods

All computations have been performed using alanine dipeptide ($AcAlaNH_2$) (see Figure 1b) as the molecular model. We have applied a B3LYP [29] density functional along with the dispersion correction (D3) [30,31] and a medium size double-zeta plus polarization basis set, SNSD [32]. The B3LYP-D3/SNSD model combines the good description of structural and spectroscopic properties of small amino-acids and dipeptides [17,33] with the empirical dispersion correction, which might be important for correct determination of the three-dimensional structure [17,18,21,22]. The B3LYP-D3 computations have been also applied for AMBER FF99SB reparameterisation allowing to deliver an improved AMBER-type force field for α,α-dialkylated-peptides [26].

All computations have been performed using GAUSSIAN suite of programs for quantum chemistry [34]. First, the relative energies along the whole Ramachandran plot were obtained from the relaxed two-dimensional (2D) scan, with the φ and ψ dihedral angles varied from -180 to 180°, in 10° intervals. For each of 1369 points in relaxed 2D scan, all other structural

parameters have been fully optimized. This provided a potential energy surface of the Ramachandran space.

The focus of this study is the “mountain-pass” regions. In order to accurately compare the QM calculated geometries with the experimental geometries, a new set of calculations were performed with the ψ and φ constrained to lie on straight lines between the energy minima and over the “mountain pass” regions. These lines define a set of the backbone Ramachandran orientations between $-60° < \varphi < +60°$ (in 5° intervals) and $\psi \approx +90°$ and −90° that follow the minimum energy configurations though both “mountain pass” regions. The slope of these lines is approximately equal to one in Ramachandran space. These leads to 50 points (25 for each of the passes) for which the local geometries were obtained by fully optimizing all other structural parameters. In order to cover all possible dielectric environments computations have been performed both in the gas phase and with $\varepsilon=78$, using polarizable continuum model [35] in its integral equation formalism (IEFPCM) implementation [36]. The resulting 100 constrained geometry optimization can be performed on a standard desktop computer. For each value of φ the structural parameters have been computed for each of the “mountain pass’ as an average of both the gas phase and polar environment. Error bars correspond to the standard deviation (SD). In order to facilitate visual comparison average parameters have been fitted to the sixth-order polynomial, reported as solid lines. The Cartesian coordinates of B3LYP-D3/SNSD/IEFPCM optimized structures corresponding to the configurations though the both “mountain pass” regions are reported in the Supporting Information. Moreover, files with Cartesian coordinates and B3LYP-D3/SNSD energies for all structures: 2D Ramachandran plot, 10° intervals, and “mountain” passes, 5° intervals are deposited at Dryad repository [37].

Results and Discussion

The computed energy contours (Fig. 2) match well Ramachandran plot [38,39], with energetically allowed large basin (−180°< φ < -60°) corresponding to the α-helix and β-sheet, a smaller one characterized by φ ≈ 60°, and a transition zone between the favored conformations. The latter span between the two regions where stable high-energy conformations (marked as blue points in Fig. 2a) have been observed [11]. It should be noted that secondary structure effects cannot be present in the potential energy surface due to the size of the peptide model.

The good agreement between computations and experimental observations validate more detailed analysis performed for the structures related to the two mountain regions marked as green lines in Figure 2a. Classical models considered these regions as disallowed due to "clashing" between the carbonyl carbon (C) and the peptide oxygen of the previous residue (O^{-1}). However, structural parameters from high-resolution structures show that in reality the observed O^{-1}…C distances are larger than 2.6 Å, with the averaged values at φ ≈ 0°, of approximately 2.7Å [11]. Figure 3 compares the computed B3LYP-D3 O^{-1}…C distances with experimentally derived averages from Ref [11] as well as standard restraints geometry and AMBER FF99SB force field values. It is clear that quantum chemical computations capture the local geometry arrangements, without any additional *ad hoc* parameterization, while the standard restraints paradigm leads to distance well below the 2.7 Å "extreme approach limit" [38] and AMBER FF99SB predicts distances too large by about 0.15 Å.

Local arrangements within O^{-1}-C^{-1}-N-Cα-C^{1} ring are also defined by the three bond angles: ∠O^{-1}-C^{-1}-N, ∠C^{-1}-N-Cα and ∠N-Cα-C^{1}, respectively. Figure 4 compares computed B3LYP-D3 values with experimentally derived averages [11] and AMBER FF99SB force field results [11]. The

experimental $\angle O^{-1}$-C^{-1}-N angle values vary between 122-124° with both FF99SB and B3LYP-D3 within 1° of the experimental results. The largest fluctuations in the reference data are observed for the $\angle C^{-1}$-N-Cα angle: experimental averages vary from 120 to 127° with the maximum for φ ≈ 0° conformations. The B3LYP-D3 results are in closer agreement with experiment than the AMBER values for all regions defined by higher data concentration (blue points). Furthermore, the QM values only slightly overestimate the average values close to maximum. By contrast, the AMBER FF99SB force field computations overestimate experimental values by 2-4° over the whole $-60° < \varphi < +60°$ range. Too large $\angle C^{-1}$-N-Cα angle is directly related to the overestimated O^{-1}…C distance. The final angle in Figure 4, the $\angle$N-Cα-C^{1} , has larger error bars due to the more scattered experimental data. The agreement between the experimental values and both the computational methods is similar over most of the range but AMBER is doing generally better near φ=0. It can be also observed that B3LYP-D3 results computed for $\psi>0°$ and $\psi<0°$ for both $\angle C^{-1}$-N-Cα and $\angle$N-Cα-C^{1} angles tend to diverge at smallest and largest values of φ. B3LYP-D3 angles are smaller for $\psi>0°$ then for $\psi<0°$ in $-60° < \varphi < -35°$ range and larger in $35° < \varphi < 60°$ range. At variance, the experimental averages have been derived from data within $-35° < \varphi < +35°$ and fitting both $\psi>0°$ and $\psi<0°$ passes together. Our computations suggest that separated functions could better describe regions beyond these boundaries. Finally we note that a good agreement between B3LYP-D3 computations and experimental data has been also obtained for other bond angles, which show much smaller variations with respect to the φ dihedral, and are reported in Figure S2. In general, the quantum mechanical computations performed using the small dipeptide molecular model predicted very accurate structural parameters within the local O^{-1}-C^{-1}-N-Cα-C^{1} arrangement for unusual conformations. This further confirms the B3LYP-

D3/SNSD as an accurate and effective computational model that can be recommended also for large scale computations of extended data sets [19,39] or more advance computational spectroscopy studies [22,40,41] of amino-acids, polypetides and protein models.

Conclusions

This work adds more evidence to the growing data that B3LYP-D3 computations can be considered as reliable and feasible tool for derivation of accurate structural parameters for non-typical Ramachandran orientations. The method can be applied to the parameterization of areas of the Ramachandran space where paucity of accurate reference experimental data would otherwise preclude any reasonable parameterization. This could be applied to the special cases of peptide links involving challenging peptides such as proline and to the rare cis-peptide or very rare cis-non-proline configurations. There are even cases of two prolines in sequence that need more reference structural information in order to provide reasonable geometries.

Moreover, small but well defined model systems and backbone orientations represent a computational strategy free from the influence of specific effects and allow determination of generic properties even for cases where their extraction from experimental data can be difficult due to lack of results spanning sufficiently different overall contexts.

Acknowledgements

Shanghai University and International Centre for Quantum and Molecular Structures are gratefully acknowledged for support. Authors thank Prof. Pavel V. Afonine and Prof. Mark P. Waller for fruitful discussions.

References

[1] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Research* 2000; **28:** 235-242; http://www.rcsb.org

[2] Balaji GA, Nagendra HG, Balaji VN, Rao SN. Experimental conformational energy maps of proteins and peptides. *Proteins* 2017; **85**:979-1001

[3] Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst.* 2010; **D66**:12-21.

[4] Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB 3rd, Snoeyink J, Richardson JS, Richardson DC. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research* 2007; **35**:W375-W383.

[5] Afonine PV, Echols N, Grosse-Kunstleve RW, Moriarty NW, Adams PD. Improved target weight optimization in phenix.refine. *Computational Crystallography Newsletter* 2011; **2**:99–103.

[6] Evans PR. An Introduction to Stereochemical Restraints. *Acta Cryst. D*: 2007; **63**:58–61.

[7] Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN. REFMAC5 Dictionary: Organization of Prior Chemical Knowledge and Guidelines for Its Use. *Acta Cryst. D* 2004; **60**:2184–95.

[8] Berkholz DS, Shapovalov MV, Dunbrack RL Jr, Karplus PA. Conformation dependence of backbone geometry in proteins. *Structur*e 2009; **17**: 1316–1325

[9] Moriarty NW, Tronrud DE, Adams PD, Karplus PA. "A New Default Restraint Library for the Protein Backbone in Phenix: A Conformation-Dependent Geometry Goes Mainstream." *Acta Cryst.. D* 2016; 72: 176–79.

[10] Moriarty NW, Tronrud DE, Adams PD, Karplus PA. Conformation-dependent backbone geometry restraints set a new standard for protein crystallographic refinement, *FEBS* 2014; **281**: 4061-71.

[11] Brereton AE, Karplus PA. Native proteins trap high-energy transit conformations. *Sci Adv* 2015; **1**: e1501188.

[12] Case DA, Darden TA, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman, AMBER 12 (2012); http://ambermd.org/

[13] Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* 2006; **65**: 712–725.

[14] Cerutti DS, Freddolino PL, Duke Jr. RE, Case DA. Simulations of a protein crystal with a high resolution x-ray structure: Evaluation of force fields and water models. *J Phys Chem B* 2010; **114**:12811–12824.

[15] Vitalini F, Mey A S J S, Noé F, Keller BG. Dynamic properties of force fields. *J Chem Phys* 2015; **142**:8.

[16] Puzzarini C, Biczysko M, Barone V, Largo L, Peña I, Cabezas C, Alonso JL. Accurate Characterization of the Peptide Linkage in the Gas Phase: A Joint Quantum-Chemical and Rotational Spectroscopy Study of the Glycine Dipeptide Analogue. *J. Phys. Chem. Lett.* 2014; **5**:534−540.

[17] Alata I, Pérez-Mellor A, Ben Nasr F, Scuderi D, Steinmetz V, Gobert F, Jaïdane N E, Zehnacker-Rentien A, Does the Residues Chirality Modify the Conformation of a Cyclo-Dipeptide? Vibrational Spectroscopy of Protonated Cyclo-diphenylalanine in the Gas Phase, *J. Phys. Chem. A* 2017; **121**:7130-7138.

[18] Stamm A, Bernhard D, Gerhards M, Structural investigations on a linear isolated depsipeptide: the importance of dispersion interactions, *Phys. Chem. Chem. Phys.* 2016; **18:**15327-15336.

[19] Ropo M, Schneider M, Baldauf C, Blum V. First-principles data set of 45,892 isolated and cation-coordinated conformers of 20 proteinogenic amino acids. *Sci Data* 2016; **3**:160009.

[20] Hostaš J, Jakubec D, Laskowski RA, Gnanasekaran R, Řezáč J, Vondrášek J, Hobza P. Representative Amino Acid Side-Chain Interactions in Protein–DNA Complexes: A Comparison of Highly Accurate Correlated Ab Initio Quantum Mechanical Calculations and Efficient Approaches for Applications to Large Systems. *J Chem Theory Comp* 2015; **11**:4086-4092.

[21] Grimme S, Hansen A, Brandenburg JG, Bannwarth C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem Rev* 2016; **116**: 5105-5154.

[22] Goerigk L, Reimers JR. Efficient Methods for the Quantum Chemical Treatment of Protein Structures: The Effects of London-Dispersion and Basis-Set Incompleteness on Peptide and Water-Cluster Geometries. *J Chem Theory Comp* 2013; **9**:3240-3251.

[23] Baldauf C, Rossi M, Going clean: structure and dynamics of peptides in the gas phase and paths to solvation. *Journal of Physics: Condensed Matter*, 2015; **27**:493002.

[24] Sousa SF, Ribeiro AJM., Neves RPP, Brás NF, Cerqueira NMFSA, Fernandes PA, Ramos MJ. Application of quantum mechanics/molecular mechanics methods in the study of enzymatic reaction mechanisms. *WIREs Comput Mol Sci*, 2017; **7**: e1281.

[25] Robertson MJ, Tirado-Rives J, Jorgensen WL. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. *J. Chem. Theory Comput.* 2015, **11**:3499–3509.

[26] Grubisic S, Brancato G, Barone V. An improved AMBER force field for α,α-dialkylated peptides: intrinsic and solvent-induced conformational preferences of model systems, *Phys. Chem. Chem. Phys.* 2013; **15**:17395-17407.

[27] Zheng M, Reimers JR, Waller MP, Afonine PV. Q|R: quantum-based refinement. *Acta Cryst.* 2017; **D73**:45-52.

[28] Zheng M, Moriarty NW, Xu Y, Reimers JR, Afonine PV, Waller MP. Solving the scalability issue in quantum-based refinement: Q|R#1. *Acta Cryst. 2017;* D73:1020-1028.

[29] Becke AD. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J Chem Phys* 1993; **98**:5648−5652.

[30] Grimme S, Antony J, Ehrlich S, Krieg H. A Consistent and Accurate ab initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J Chem Phys* 2010; **132**:154104.

[31] Grimme S, Ehrlich S, Goerigk L, Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem*. 2011; **32**:1456-1465.

[32] Barone V, Biczysko M, Bloino J. Fully Anharmonic, IR and Raman Spectra of Medium-size Molecular Systems: Accuracy and Interpretation. *Phys Chem Chem Phys* 2014; **16**:1759−1787.

[33] Barone V, Biczysko M, Bloino J, Puzzarini C. The Performance of Composite Schemes and Hybrid CC/DFT Model in Predicting Structure, Thermodynamic and Spectroscopic Parameters: the Challenge of the Conformational Equilibrium in Glycine. *Phys. Chem. Chem. Phys.* 2013; **15**: 10094-10111.

[34] Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery Jr. JA, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Keith T, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross BC, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox D, Gaussian 09, Revision E.01, Gaussian, Inc., Wallingford, CT, 2015

[35] Tomasi J, Mennucci B, Cammi R. Quantum mechanical continuum solvation models. *Chem Rev* 2005; **105**:2999-3093.

[36] Scalmani G, Frisch MJ. Continuous surface charge polarizable continuum models of solvation. I. General formalism. *J Chem Phys* 2010; **132**:114110

[37] Jiang Z, Biczysko M, Moriarty NW. Data from: Accurate geometries for “mountain pass” regions of the Ramachandran plot using quantum chemical calculations. DOI: https://doi.org/10.5061/dryad.hkxxxx

[38] Ramachandran GN, Sasisekharan V, Conformation of polypeptides and proteins. *Adv Protein Chem* 1968; **23**:283–438.

[39] Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM. Prisant MG, Richardson JS, Richardson DC. Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. *Proteins* 2003; **50**: 437–450.

[40] Najbauer EE, Bazso G, Apostolo R, Fausto R, Biczysko M, Barone V, Tarczay G. Identification of Serine Conformers by Matrix-Isolation IR Spectroscopy Aided by Near-Infrared Laser-Induced Conformational Change, 2D Correlation Analysis, and Quantum Mechanical Anharmonic Computations. *J Phys Chem B* 2015; **119**:10496-10510

[41] Mahe J, Jaeqx S, Rijs AM, Gaigeot MP, Can far-IR action spectroscopy combined with BOMD simulations be conformation selective? *Phys Chem Chem Phys* 2015; **17**:25905-25914.

Figure Legends

Figure 1. Schematic representation of peptide bonding for “mountain region” Ala (A:12) residue from PDB entry 3VUR (a) and $AcAlaNH_2$ molecular model (b).

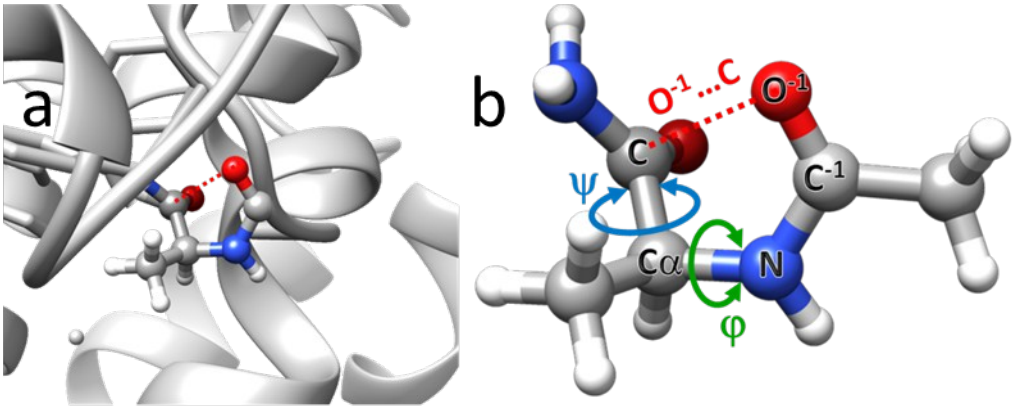
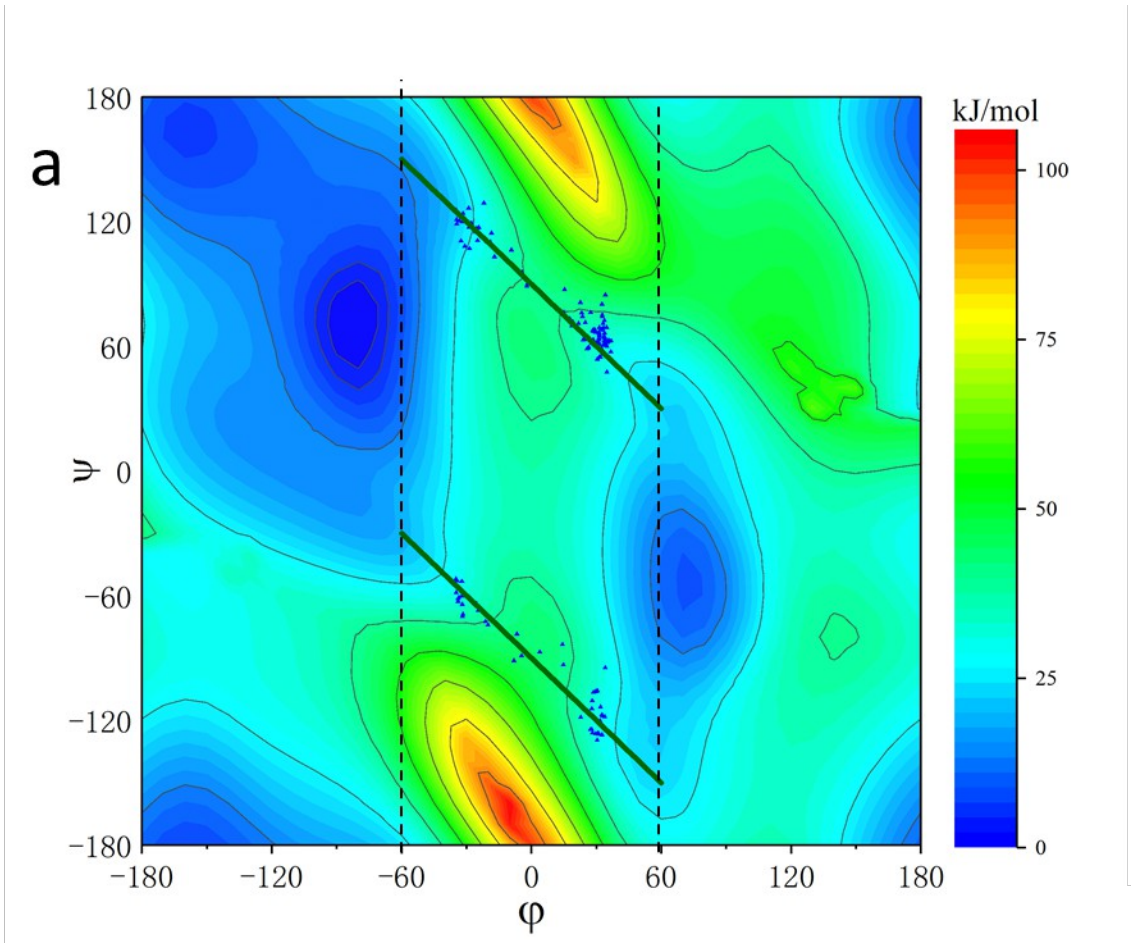


Figure 2. Energy contours for $AcAlaNH_2$ (from relaxed scan computed at the B3LYP-D3/SNSD level) along with (a) experimental data between $-35° < \varphi < +35°$ from Ref. [11], all 146 reliable observations (blue points) and the best-fit lines (green, see text and Ref. [11] for the details); (b) plot showing about 100,000 data points (black dots) for general amino-acid types (not Gly, Pro, or pre-Pro) in high-resolution crystal structures from Ref [39].


a
180
120
60
ψ 0
−60
−120
−180
−180
−120
−60
0
60
120
180
φ
kJ/mol
100
75
50
25
0

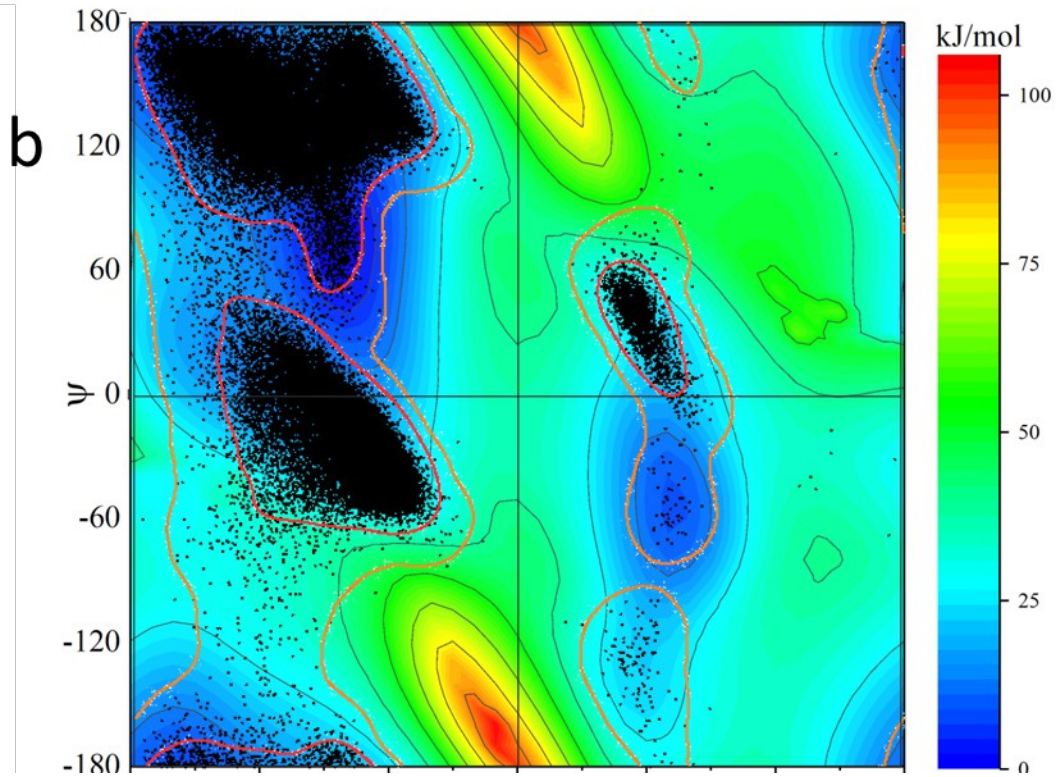

b
180
120
60
ψ 0
-60
-120
-180
kJ/mol
100
75
50
25
0

Figure 3. O^{-1}…C distance plotted as a function of φ: B3LYP-D3/SNSD restrained optimization (B3D3, ψ>0° purple, ψ<0° pink, averaged between two calculations, the gas phase and ε=78 dielectric environment) compared to data from Ref. [11]: observed average (PDB, green dots and error bars), each data point from observations between −35° < φ < +35° (blue points), empirically defined φ-dependent geometry functions (green line), standard geometry (StG, black line), the AMBER FF99SB force field (FF99SB, orange line).

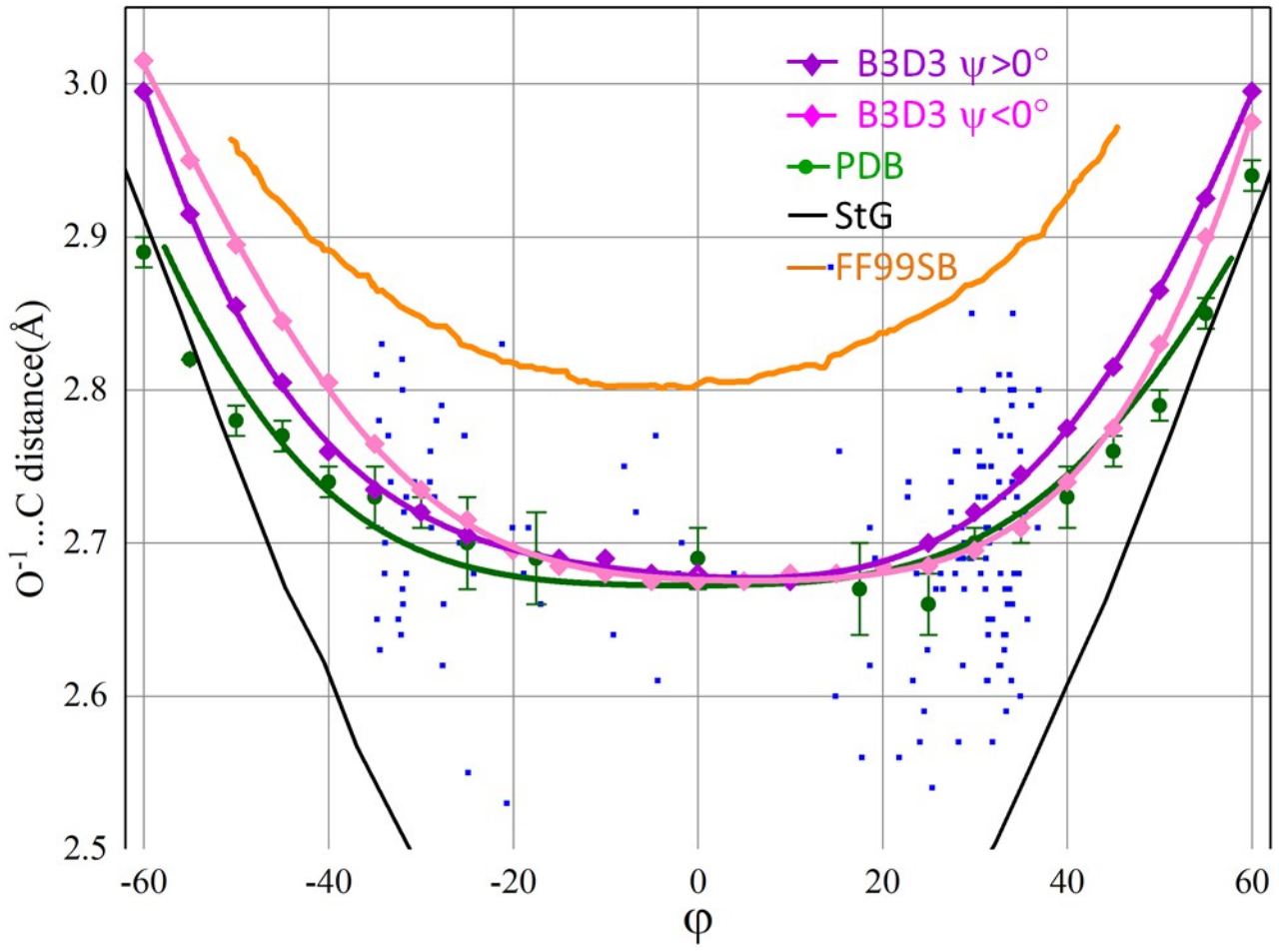
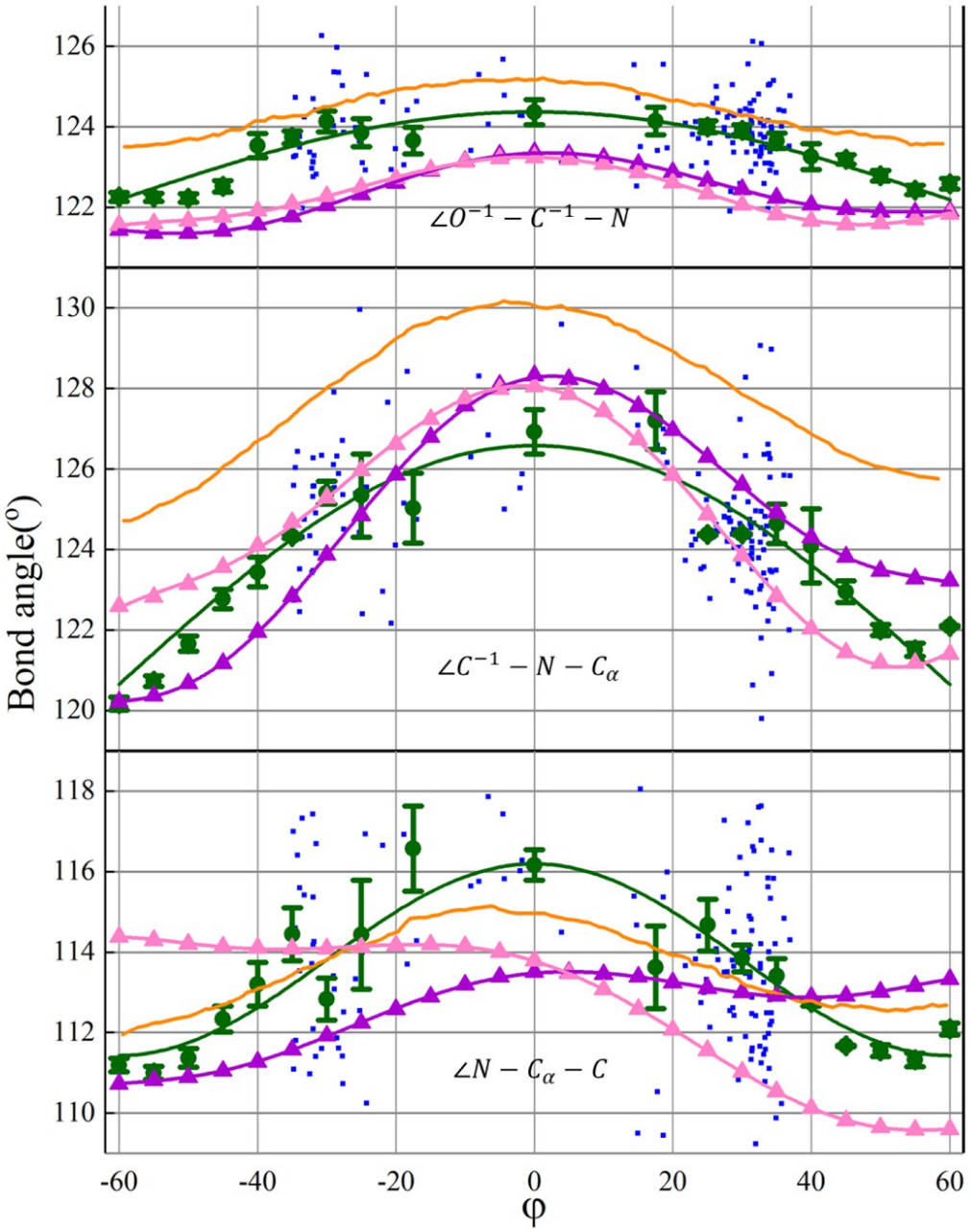


Figure 4. $\angle O^{-1}\text{-}C^{-1}\text{-}N$, $\angle C^{-1}\text{-}N\text{-}C\alpha$ and $\angle N\text{-}C\alpha\text{-}C^{1}$, bond angles plotted as a function of φ: B3LYP-D3/SNSD restrained optimization (ψ>0° purple, ψ<0° pink, averaged between two calculations, the gas phase and ε=78 dielectric environment) compared to data from Ref. [11]: observed average (PDB, green dots and error bars), each data point from observations between $-35° < \varphi < +35°$ (blue points), empirically defined φ-dependent geometry functions (green line), the AMBER FF99SB force field (FF99SB, orange line).


126
124
122
$\angle O^{-1} - C^{-1} - N$
130
128
126
124
122
120
$\angle C^{-1} - N - C_\alpha$
118
116
114
112
110
$\angle N - C_\alpha - C$
Bond angle(°)
-60
-40
-20
0
20
40
60
φ