

A Theoretical Investigation of Roles of Backbone Amide Resonance in Protein Structure

John Neville Sharley

Discipline of Chemistry, School of Physical Sciences, University of Adelaide

Submitted to the University of Adelaide in April 2016 for the degree of Doctor of Philosophy

Abstract 5

Declaration 7

Acknowledgements 8

1 Introduction 9

- 1.1 *Classical view 9*
- 1.2 *Current view 10*
- 1.3 *Backbone-based view 11*

2 Natural Bond Orbital analysis 14

- 2.1 *Suitability for bonding analysis 14*
- 2.2 *Output in standard format 14*
- 2.3 *Querying XML 16*
- 2.4 *Automated construction of experiments 18*

3 Established DFT methods calculation of conjugation disturbed in the presence of torsional hyperconjugation 20

- 3.1 *Statement of authorship 20*
- 3.2 *Author contact 20*
- 3.3 *Abstract 21*
- 3.4 *Notation 21*
- 3.5 *Overview 21*
- 3.6 *Introduction 23*
- 3.7 *Methods 24*
- 3.8 *Results and discussion 24*
 - 3.8.1 *Vinylamine 24*
 - 3.8.2 *Ethanamide 31*
 - 3.8.2.1 *Carbonyl orbital disturbances 31*
 - 3.8.2.2 *Pyramidalization at nitrogen 34*
 - 3.8.3 *Antiparallel beta sheet 39*
- 3.9 *Conclusion 40*
- 3.10 *Acknowledgements 42*
- 3.11 *Appendix 1 43*
- 3.12 *Appendix 2 72*

4 Variation of protein backbone amide resonance by electrostatic field 75

- 4.1 *Statement of authorship 75*
- 4.2 *Author contact 75*
- 4.3 *Abstract 76*
- 4.4 *Key phrases 76*
- 4.5 *Notation 77*
- 4.6 *Overview 77*
- 4.7 *Introduction 78*
- 4.8 *Software 79*
- 4.9 *Results and discussion 79*
 - 4.9.1 *Constrained to O-C-N plane 79*
 - 4.9.2 *Partial sp³ hybridization at N 85*
 - 4.9.3 *Other geometry variations 87*
 - 4.9.4 *Torsional steering 87*
 - 4.9.5 *Backbone amide nitrogen as hydrogen bond acceptor 89*
 - 4.9.6 *RAHB chains in electrostatic field 90*
 - 4.9.7 *Electrostatic field vectors at backbone amides 93*
 - 4.9.8 *Protein beta sheet 94*
 - 4.9.9 *RAHB protein helices 95*

- 4.9.10 Proline 96
- 4.9.11 Permittivity 96
- 4.9.12 Protein folding, conformational and allosteric change 97
- 4.9.13 Unstructured regions 97
- 4.9.14 Sources of electrostatic field variation in proteins 98
- 4.9.15 Absence of charged residues 98
- 4.9.16 Development of methods 98
- 4.10 *Hypotheses* 99
 - 4.10.1 Beta sheet 99
 - 4.10.2 Amyloid fibril 100
 - 4.10.3 Polyproline helices types I and II 101
 - 4.10.4 Protein folding 102
 - 4.10.4.1 Overview 102
 - 4.10.4.2 Peptide group resonance drives protein folding 103
 - 4.10.4.3 Specification of fold separated from how to fold 104
 - 4.10.4.4 Comparison with earlier backbone-based theory 104
 - 4.10.4.5 Observation 106
 - 4.10.4.6 Summary 107
 - 4.10.5 Molecular chaperones and protein complexes 107
 - 4.10.6 Nitrogenous base pairing 108
- 4.11 *Conclusion* 110
- 4.12 *Acknowledgements* 111
- 4.13 *Appendix 1* 112
- 4.14 *Appendix 2* 132

5 Do cooperative cycles of hydrogen bonding exist in proteins?

133

- 5.1 *Statement of authorship* 133
- 5.2 *Author contact* 133
- 5.3 *Abstract* 134
- 5.4 *Introduction* 134
 - 5.4.1 Review of Resonance-Assisted HB 134
 - 5.4.2 Phenomena anticipated if cyclic cooperative HB exists in proteins 135
- 5.5 *Notation* 136
- 5.6 *Methods* 136
 - 5.6.1 Counterpoise correction 136
 - 5.6.2 Dispersion correction 137
 - 5.6.3 Software packages 137
 - 5.6.4 Haskell 137
- 5.7 *Results and Discussion* 138
 - 5.7.1 HB angle 138
 - 5.7.1.1 Experiments involving hydrogen bonding between an amide group and hydrogen fluoride 138
 - 5.7.1.2 Use of hydrogen bonding to carbonyl oxygen where C-O-N is far from linear 144
 - 5.7.1.3 Experiments involving hydrogen bonding between two amide groups 145
 - 5.7.2 Inter-backbone amide hydrogen bonding C-O-N angles in the Protein Data Bank 148
 - 5.7.3 Linear chain of formamides 148
 - 5.7.4 Formamide cycles 150
 - 5.7.5 Three alpha helices 153
 - 5.7.6 Artificial beta sheets 156
 - 5.7.7 Cyclic HB in the Protein Data Bank 157
 - 5.7.8 Joined HB chains of beta sheet 162
- 5.8 *Conclusion* 165
- 5.9 *Acknowledgements* 168
- 5.10 *Appendix 1* 169
- 5.11 *Appendix 2. Polyvaline Parallel Beta Sheet* 183
- 5.12 *Appendix 3. Polyvaline Antiparallel Beta Sheet* 187

6 Amino acid preference against beta sheet through allowing backbone hydration enabled by the presence of cation 191

- 6.1 *Statement of authorship 191*
- 6.2 *Author contact 191*
- 6.3 *Abstract 192*
- 6.4 *Introduction 192*
 - 6.4.1 *Alpha helix preferring amino acid residues in a beta sheet 192*
 - 6.4.2 *Cation interactions with protein backbone oxygen 193*
 - 6.4.3 *Quantum molecular dynamics with quantum mechanical treatment of every water molecule 193*
- 6.5 *Methods 194*
- 6.6 *Results 195*
 - 6.6.1 *Preparation 195*
 - 6.6.2 *Experiment 1302 195*
 - 6.6.3 *Experiment 1303 197*
- 6.7 *Discussion 199*
 - 6.7.1 *HB networks of water 199*
 - 6.7.2 *Subsequent to rupture of a transient beta sheet 199*
 - 6.7.3 *Hofmeister effects 200*
- 6.8 *Conclusion 201*
- 6.9 *Future work 202*
- 6.10 *Acknowledgements 203*
- 6.11 *Appendix 1. Backbone hydration in experiment 1302 204*
- 6.12 *Appendix 2. Backbone hydration in experiment 1303 205*

7 Conclusion 207

- 7.1 *Significance and contribution 207*
- 7.2 *Future work 210*

References 211

Abstract

Protein structure, including its dynamics, is of pervasive significance in biology. A protein's structure determines its bindings with other molecules [1], and from that its roles in signal transduction, enzymatic activity and mechanical structure. Few cellular processes have no protein involvement. The relationship between a protein's sequence of amino-acid residues and its three-dimensional structure, partial or otherwise, has long been of considerable interest [2].

A theory of protein folding is proposed in Section 4.10.4 (Hypotheses/Protein folding). This theory varies and extends the backbone-based theory proposed by Rose *et al.* [3]. This theory may prove to be the most significant offering of the thesis.

Overall, this thesis investigates the variation in peptide group resonance and its implications for Resonance-Assisted Hydrogen Bonding [4, 5], RAHB, such as exists in inter-peptide group hydrogen bonding. Natural Bond Orbital [6], NBO, analysis is used for this investigation, and Section 2.1 summarizes the virtues of NBO.

Chapter 3 is concerned with methods for computational investigation of protein structures, and finds that methods and basis sets most often used for these investigations are particularly unsuitable when any beta sheets are present due to poor modelling of amide resonance and hence of RAHB that features in the hydrogen-bonded chains of backbone amides of protein secondary structures such as beta sheets [7] and alpha helices [8].

Chapter 4 reports experiments quantifying the sensitivity of amide resonance to electrostatic field with component parallel or antiparallel to the amide C-N bond. This sensitivity allows electrostatic properties including permittivity of amino-acid residue sidechains to influence backbone amide resonance and hence secondary structure RAHB chains, giving a novel mechanism relating residue sequence to structure. Also, this variation in amide resonance is energetically significant even without considering a hydrogen bonding context. Variation of peptide group resonance is expected to vary the barrier height of prolyl isomerization [9]. Subsequent to quantifying this effect in isolated amides and in a RAHB chain, hypotheses are offered concerning the stability of beta sheets, amyloid fibrils [10] and polyproline helices [11]. An analogous sensitivity in nitrogenous base pairing [12] is conjectured. A hypothesis is offered concerning protein complexation and molecular chaperone [13] action.

Chapter 5 is motivated by the observed increase of stabilization when cooperative hydrogen bonding chains are cyclized in non-protein contexts [6] and by the phenomena anticipated if these cycles exist in proteins as described in the chapter. The question of the optimal geometry for amide-amide hydrogen bonding is revisited with emphasis on the inequivalence of amide oxygen lone pairs. A possible avenue for the design of HB-chain polymers with improved stability is discussed.

Chapter 6 studies a dependency of amino acid residue preference against beta sheet secondary structure by backbone hydration in the presence of cation, doing so by Quantum Molecular Dynamic simulation of a beta sheet with a full quantum mechanical treatment of each solvent molecule.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Name: John Neville Sharley

Signed:

Dated: 2016-09-26

Acknowledgements

Prof. John A. Carver is acknowledged for reading this thesis and offering editing suggestions. I am profoundly grateful for the enlightened research environment he provided.

Prof. Peter M. W. Gill is acknowledged for reading the earliest draft of Chapter 3 and pointing out the resemblance of the second mentioned problem, pyramidalization at amide nitrogen at wavefunction method/Pople basis set, to the benzene non-planarity problem [14].

I claim that there is ample evidence here that pre-candidature mentoring by Iain Murchland and Cvetan Stojkoski resulted in public good.

eResearch South Australia is acknowledged for hosting and administering machines provided under Australian Government Linkage, Infrastructure, Equipment and Facilities grants for Supercomputing in South Australia, directing funds to the acquisition of Nvidia Tesla GPU nodes and allocating 64 CPU cores and 256 GB RAM of the NeCTAR Research Cloud (a collaborative Australian research platform supported by the National Collaborative Research Infrastructure Strategy) to the present work.