Analysis of interactions between secondary structures in globular proteins
by Robert Barton Burris

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biochemistry
Montana State University
© Copyright by Robert Barton Burris (1992)

Abstract:
In an attempt to develop a system to aid protein structure prediction and the design of de novo proteins, residue-residue interactions in globular proteins are analyzed. Computer programs were written which studied interactions between charged atoms, polar side-chain atoms, polar backbone atoms, total backbone atoms and hydrophobic side-chain atoms. The data from these studies are analyzed and improvements in programs and new avenues of investigation are suggested.

Programs were written which would evaluate the data from the above programs. These programs have been designed to analyze the data from several perspectives. Interactions between residues were studied for the entire protein, between secondary structures in general, and between the following combinations of secondary structure pairs: 1)α-helix - α-helix, 2)α-helix - β-sheet, 3)β-sheet - α-helix, 4)β-sheet - β-sheet. For interactions between hydrophobic side chain atoms, the data were analyzed further. Interactions between pairs of residues were studied according to the angles at which the secondary structures crossed. The number of occurrences of each amino acid at each position relative to the nearest backbone contact are presented for all angle ranges. Additionally, the number of residue-residue contacts and average number of atom contacts are given for each amino acid at each position. To demonstrate the capabilities of the system to analyze data, Ala was studied for interactions between pairs of α-helices with crossing angles of -150° to -180°. For Ala at each position relative to nearest backbone atom contact, all atoms which interacted with Ala were listed according to their relative position to the nearest backbone atom contact, the number of residue-residue interactions the residue made with Ala, and the average number of atom contacts made between the two residues.

The results of the above analyses were evaluated for the heat stable protein thermolysin. The analysis of this thermophilic protein was undertaken in an attempt to provide insight into protein stability with the ultimate goal of assisting in the design of novel proteins. It is suggested that triads of α-helices, all of which interact with each other, contribute to the stability of thermolysin.

ANALYSIS OF INTERACTIONS BETWEEN SECONDARY STRUCTURES IN GLOBULAR PROTEINS

by

ROBERT BARTON BURRIS

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Biochemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

November, 1992

D378
B9446

APPROVAL

of a thesis submitted by

Robert Barton Burris

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

11/24/92
Date

Samuel J. Rogers
Chairperson, Graduate Committee

Approved for the Major Department

11/24/92
Date

John R. Amend
Head, Major Department

Approved for the College of Graduate Studies

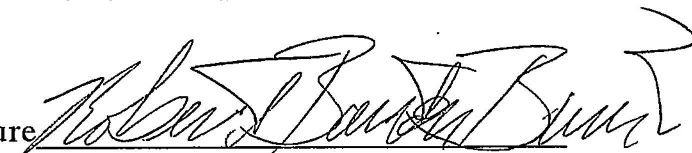
1/11/93
Date

R L Brown
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or in parts may be granted only by the copyright holder.

Signature 

Date 12/29/92

TABLE OF CONTENTS

Page

LIST OF TABLES viii

ABSTRACT xxxi

INTRODUCTION 1

- Protein Folding 1
  - Energy of Folding 1
  - Charged Groups 3
  - Hydrogen Bonds 8
- Secondary Structures 9
  - α-Helices 10
  - ß-Sheets 11
- Motifs 12
  - Beta-Alpha-Beta Motifs 12
  - Beta-Turn-Beta 13
  - Helix-Turn-Helix 13
- Side-Chain Packing 13
  - α-helices 13
  - Helix-Sheet 15
  - Sheet-Sheet 16
- Computer Programs 16
  - Predictive Programs 16
  - Analytical Programs 19

STATEMENT OF THE PROBLEM 22

EXPERIMENTAL 24

- Selection of Database 24
  - Strategy 24
- Core Programs 27
  - Backbone - Backbone Contacts 27
  - Backbone - Hydrophobic Side Chain Contacts 28
  - Hydrophobic Side Chain - Hydrophobic Side Chain Contacts 28
  - Charged - Polar Side Chain Contacts 28

Page

Charged - Backbone Carbonyl Oxygen Contacts 28
Charged - Backbone Amide Nitrogen Contacts 28
Polar Side Chain - Polar Side Chain Contacts 30
Polar Side Chain - Backbone Carbonyl Oxygen Contacts 30
Polar Side Chain - Backbone Amide Nitrogen Contacts 30
Oppositely Charged Atom Contacts 30
Like Charged Atom Contacts 30
Angle Determination 31
Angle 32
Nearest-Backbone-Contact 35
Angle-Nearest-Backbone-Combine 35
Angle-Sign Determination 35
Programs Which Combine Data 38
Combine 38
Secondary Structure-Total-Combine 41
Range 41
Total-Residue Pair-Interactions 41
Secondary Structure-Residue Pair-Interactions 41
Helix-Sheet-Residue Pair-Interactions 46
Distribution-Count 46
Distribution-Sum 46
Residue-Position-Contacts 47
Residue-Occurrence-By Position 47
Protein Profile 47
Secondary Structure Profile 48
Total-Atom Contacts 48
Secondary Structure- Atom Contacts 48
Total-Residue Contacts 48
Secondary Structure-Residue Contacts 48
Total-Residue Occurrence 48
Secondary Structure-Residue Occurrence 49
Total-Contact Ratio 49
Secondary Structure-Contact Ratio 50
Angle Count 50

RESULTS AND DISCUSSION 51
Protein Interactions 51
Interactions Between Oppositely Charged Groups 52
Interactions Between Residues of Like Charge 52

Page

Interactions Between Charged Atoms and Polar Backbone Atoms 53
Charged Atom - Polar Side Chain Atom Interactions 54
Polar Side Chain Atom - Polar Side Chain Atom Interactions 55
Polar Side Chain Atom - Polar Backbone Atom Interactions 56
Backbone Atom - Backbone Atom Interactions 56
Backbone Atom - Hydrophobic Side Chain Atom Interactions 57
Hydrophobic Side Chain - Hydrophobic Side Chain Interactions 58
Hydrophobic Side Chain Interactions by Angle Ranges 59
Helix-Helix 60
Helix-Sheet 61
Sheet-Sheet 62
Ala Interactions in Helical Pairs with Crossing Angles between -150° and -180° 62
Interactions by Position 64
Backbone-Backbone Interactions 64
Backbone-Hydrophobic Side Chain Interactions 65
Charged Atom-Polar Backbone Atom Interactions 65
Charged Atom-Polar Side Chain Interactions 66
Hydrophobic Side Chain Atom Interactions 66
Charged Atom Interactions 66
Polar Side Chain-Polar Backbone Interactions 67
Polar Side Chain Interactions 67
Preferences by Position 67
Helix-Helix Interactions 68
Helix-Sheet Interactions 69
Sheet-Sheet Interactions 70
Contact Ratios 71
Thermolysin 73
Interactions between Secondary Structures 73

SUMMARY 77

REFERENCES CITED 81

APPENDICES 85

Page

Appendix A 86
Appendix B 92
Appendix C 100
Appendix D 113
Appendix E 120
Appendix F 127
Appendix G 140
Appendix H 153
Appendix I 160
Appendix J 197
Appendix K 246
Appendix L 288
Appendix M 296

LIST OF TABLES

Table **Page**

1. Sample output from the program Hydrophobic Side-Chain - Hydrophobic Side-Chain Contacts for Thermolysin 29

2. Sample output from the program Combine for thermolysin 39

3. Sample output from the program Range for thermolysin 42

4. Sample output from the program Secondary Structure Total Combine for thermolysin 44

5. Total number of residue-residue interactions and average atom contacts for interactions between oppositely charged atoms 87

6. Total number of residue-residue interactions and average atom contacts for interactions between secondary structures for oppositely charged atoms 88

7. Number of residue-residue interactions and average atom contacts for interactions between oppositely charged atoms. Interactions are between pairs of α-helices 89

8. Number of residue-residue interactions and average atom contacts for interactions between oppositely charged atoms. Interactions are between α-helices and ß-sheets 90

9. Number of residue-residue interactions and average atom contacts for interactions between oppositely charged atoms. Interactions are between ß-sheets and α-helices 91

10. Number of residue-residue interactions and average atom contacts for interactions between oppositely charged atoms. Interactions are between pairs of ß-sheets 93

LIST OF TABLES-continued

Table Page

11. Total number of residue-residue interactions and average atom contacts for interactions between atoms of like charge ... 94

12. Total number of residue-residue interactions and average atom contacts for interactions between secondary structures for atoms of like charge ... 95

13. Number of residue-residue interactions and average atom contacts for interactions between atoms of like charge. Interactions are between pairs of α-helices ... 96

14. Number of residue-residue interactions and average atom contacts for interactions between atoms of like charge. Interactions are between α-helices and ß-sheets ... 97

15. Number of residue-residue interactions and average atom contacts for interactions between atoms of like charge. Interactions are between ß-sheets and α-helices ... 98

16. Number of residue-residue interactions and average atom contacts for interactions between atoms of like charge. Interactions are between pairs of ß-sheets ... 99

17. Total number of residue-residue interactions and average atom contacts for interactions between charged atoms and backbone carbonyl oxygen atoms ... 101

18. Total number of residue-residue interactions and average atom contacts for interactions between secondary structures for charged atoms and backbone carbonyl atoms ... 102

19. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and backbone carbonyl oxygen atoms. Interactions are between pairs of α-helices ... 103

LIST OF TABLES-continued

Table Page

20. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and backbone carbonyl oxygen atoms. Interactions are between α-helices and ß-sheets 104

21. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and backbone carbonyl oxygen atoms. Interactions are between ß-sheets and α-helices 105

22. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and backbone carbonyl oxygen atoms. Interactions are between pairs of ß-sheets 106

23. Total number of residue-residue interactions and average atom contacts for interactions between charged atoms and backbone nitrogen atoms 107

24. Total number of residue-residue interactions and average atom contacts for interactions between secondary structures for charged atoms and backbone nitrogen atoms 108

25. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and backbone nitrogen atoms. Interactions are between pairs of α-helices 109

26. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and backbone nitrogen atoms. Interactions are between α-helices and ß-sheets 110

27. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and backbone nitrogen atoms. Interactions are between ß-sheets and α-helices 111

LIST OF TABLES-continued

Table Page

28. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and backbone nitrogen atoms. Interactions are between pairs of ß-sheets 112

29. Total number of residue-residue interactions and average atom contacts for interactions between charged atoms and polar side-chain atoms 114

30. Total number of residue-residue interactions and average atom contacts for interactions between secondary structures for charged atoms and polar side-chain atoms 115

31. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and polar side-chain atoms. Interactions are between pairs of α-helices 116

32. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and polar side-chain atoms. Interactions are between α-helices and ß-sheets 117

33. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and polar side-chain atoms. Interactions are between ß-sheets and α-helices 118

34. Number of residue-residue interactions and average atom contacts for interactions between charged atoms and polar side-chain atoms. Interactions are between pairs of ß-sheets 119

35. Total number of residue-residue interactions and average atom contacts for interactions between polar side-chain atoms 121

36. Total number of residue-residue interactions and average atom contacts for interactions between secondary structures for interactions between polar side-chain atoms 122

LIST OF TABLES-continued

Table Page

37. Number of residue-residue interactions and average atom contacts between polar side-chain atoms. Interactions are between pairs of α-helices ... 123

38. Number of residue-residue interactions and average atom contacts between polar side-chain atoms. Interactions are between α-helices and ß-sheets ... 124

39. Number of residue-residue interactions and average atom contacts between polar side-chain atoms. Interactions are between ß-sheets and α-helices ... 125

40. Number of residue-residue interactions and average atom contacts between polar side-chain atoms. Interactions are between pairs of ß-sheets ... 126

41. Total number of residue-residue interactions and average atom contacts between polar side-chain atoms backbone oxygen atoms ... 128

42. Total number of residue-residue interactions and average atom contacts between secondary structures for polar side-chain atoms and backbone oxygen atoms ... 129

43. Number of residue-residue interactions and average number of atom contacts between polar side-chain atoms and backbone oxygen atoms. Interactions are between pairs of α-helices ... 130

44. Number of residue-residue interactions and average number of atom contacts between polar side-chain atoms and backbone oxygen atoms. Interactions are between α-helices and ß-sheets ... 131

45. Number of residue-residue interactions and average number of atom contacts between polar side-chain atoms and backbone oxygen atoms. Interactions are between ß-sheets and α-helices ... 132

LIST OF TABLES-continued

Table | Page

46. Number of residue-residue interactions and average number of atom contacts between polar side-chain atoms and backbone oxygen atoms. Interactions are between pairs of ß-sheets 133

47. Total number of residue-residue interactions and average atom contacts between polar side-chain atoms backbone nitrogen atoms 134

48. Total number of residue-residue interactions and average atom contacts between secondary structures for polar side-chain atoms and backbone nitrogen atoms 135

49. Number of residue-residue interactions and average number of atom contacts between polar side-chain atoms and backbone nitrogen atoms. Interactions are between pairs of α-helices 136

50. Number of residue-residue interactions and average number of atom contacts between polar side-chain atoms and backbone nitrogen atoms. Interactions are between α-helices and ß-sheets 137

51. Number of residue-residue interactions and average number of atom contacts between polar side-chain atoms and backbone nitrogen atoms. Interactions are between ß-sheets and α-helices 138

52. Number of residue-residue interactions and average number of atom contacts between polar side-chain atoms and backbone nitrogen atoms. Interactions are between pairs of ß-sheets 139

53. Total number of residue-residue interactions and average atom contacts for interactions between backbone atoms 141

54. Total number of residue-residue interactions and average atom contacts for interactions between secondary structures for interactions between backbone atoms 142

LIST OF TABLES-continued

Table Page

55. Number of residue-residue interactions and average atom contacts between backbone atoms. Interactions are between pairs of α-helices 143

56. Number of residue-residue interactions and average atom contacts between backbone atoms. Interactions are between α-helices and ß-sheets 144

57. Number of residue-residue interactions and average atom contacts between backbone atoms. Interactions are between ß-sheets and α-helices 145

58. Number of residue-residue interactions and average atom contacts between backbone atoms. Interactions are between pairs of ß-sheets 146

59. Total number of residue-residue interactions and average atom contacts for interactions between backbone atoms and hydrophobic side-chain atoms 147

60. Total number of residue-residue interactions and average atom contacts for interactions between secondary structures for backbone atoms and hydrophobic side-chain atoms 148

61. Number of residue-residue interactions and average atom contacts between backbone atoms and hydrophobic side-chain atoms. Interactions are between pairs of α-helices 149

62. Number of residue-residue interactions and average atom contacts between backbone atoms and hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets 150

63. Number of residue-residue interactions and average atom contacts between backbone atoms and hydrophobic side-chain atoms. Interactions are between ß-sheets and α-helices 151

64. Number of residue-residue interactions and average atom contacts between backbone atoms and hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets 152

LIST OF TABLES-continued

Table | Page

65. Total number of residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms 154

66. Total number of residue-residue interactions and average atom contacts for interactions between secondary structures for interactions between hydrophobic side-chain atoms 155

67. Number of residue-residue interactions and average atom contacts between hydrophobic side-chain atoms. Interactions are between pairs of α-helices 156

68. Number of residue-residue interactions and average atom contacts between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets 157

69. Number of residue-residue interactions and average atom contacts between hydrophobic side-chain atoms. Interactions are between ß-sheets and α-helices 158

70. Number of residue-residue interactions and average atom contacts between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets 159

71. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between 0° and 30° 161

72. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between 30° and 60° 162

73. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between 60° and 90° 163

LIST OF TABLES-continued

Table | Page

74. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between 90° and 120° 164

75. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between 120° and 150° 165

76. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between 150° and 180° 166

77. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between -30° and 0° 167

78. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between -60° and -30° 168

79. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between -90° and -60° 169

80. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between -120° and -90° 170

81. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between -150° and -120° 171

LIST OF TABLES-continued

Table | Page

82. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of α-helices with crossing angles between -180° and -150° 172

83. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between 0° and 30° 173

84. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between 30° and 60° 174

85. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between 60° and 90° 175

86. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between 90° and 120° 176

87. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between 120° and 150° 177

88. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between 150° and 180° 178

89. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between -30° and 0° 179

LIST OF TABLES-continued

Table | Page

90. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between -60° and -30° 180

91. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between -90° and -60° 181

92. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between -120° and -90° 182

93. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between -150° and -120° 183

94. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between α-helices and ß-sheets with crossing angles between -180° and -150° 184

95. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between 0° and 30° 185

96. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between 30° and 60° 186

97. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between 60° and 90° 187

LIST OF TABLES-continued

Table Page

98. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between 90° and 120° 188

99. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between 120° and 150° 189

100. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between 150° and 180° 190

101. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between -30° and 0° 191

102. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between -60° and -30° 192

103. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between -90° and -60° 193

104. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between -120° and -90° 194

105. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between -150° and -120° 195

LIST OF TABLES-continued

Table | Page

106. Total residue-residue interactions and average atom contacts for interactions between hydrophobic side-chain atoms. Interactions are between pairs of ß-sheets with crossing angles between -180° and -150 196

107. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between 0° and 30° 198

108. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between 30° and 60° 199

109. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between 60° and 90° 200

110. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between 90° and 120° 201

111. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between 120° and 150° 202

112. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between 150° and 180° 203

113. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between -30° and 0° 204

114. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between -60° and -30° 205

115. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between -90° and -60° 206

LIST OF TABLES-continued

Table Page

116. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between -120° and -90° 207

117. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between -150° and -120° 208

118. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of α-helices with crossing angles between -180° and -150° 209

119. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between 0° and 30° 210

120. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between 30° and 60° 211

121. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between 60° and 90° 212

122. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between 90° and 120° 213

123. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between 120° and 150° 214

124. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between 150° and 180° 215

125. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between -30° and 0° 216

LIST OF TABLES-continued

Table | Page

126. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between -60° and -30° 217

127. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between -90° and -60° 218

128. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between -120° and -90° 219

129. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between -150° and -120° 220

130. Amino acid preferences by relative position with respect to nearest backbone contacts for α-helices and ß-sheets with crossing angles between -180° and -150° 221

131. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between 0° and 30° 222

132. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between 30° and 60° 223

133. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between 60° and 90° 224

134. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between 90° and 120° 225

135. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between 120° and 150° 226

LIST OF TABLES-continued

Table Page

136. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between 150° and 180° 227

137. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between -30° and 0° 228

138. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between -60° and -30° 229

139. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between -90° and -60° 230

140. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between -120° and -90° 231

141. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between -150° and -120° 232

142. Amino acid preferences by relative position with respect to nearest backbone contacts for ß-sheets and α-helices with crossing angles between -180° and -150° 233

143. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between 0° and 30° 234

144. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between 30° and 60° 235

145. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between 60° and 90° 236

LIST OF TABLES-continued

Table | Page

146. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between 90° and 120° 237

147. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between 120° and 150° 238

148. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between 150° and 180° 239

149. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between -30° and 0° 240

150. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between -60° and -30° 241

151. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between -90° and -60° 242

152. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between -120° and -90° 243

153. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between -150° and -120° 244

154. Amino acid preferences by relative position with respect to nearest backbone contacts for pairs of ß-sheets with crossing angles between -180° and -150° 245

155. Number of interactions between oppositely charged atoms and average number of atom contacts at each position for interactions between pairs of α-helices 247

LIST OF TABLES-continued

Table | Page

156. Number of interactions between oppositely charged atoms and average number of atom contacts at each position for interactions between α-helices and β-sheets 248

157. Number of interactions between oppositely charged atoms and average number of atom contacts at each position for interactions between β-sheets and α-helices 249

158. Number of interactions between oppositely charged atoms and average number of atom contacts at each position for interactions between pairs of β-sheets 250

159. Number of interactions between atoms of like charge and average number of atom contacts at each position for interactions between pairs of α-helices 251

160. Number of interactions between atoms of like charge and average number of atom contacts at each position for interactions between α-helices and β-sheets 252

161. Number of interactions between atoms of like charge and average number of atom contacts at each position for interactions between β-sheets and α-helices 253

162. Number of interactions between atoms of like charge and average number of atom contacts at each position for interactions between pairs of β-sheets 254

163. Number of interactions between charged atoms and polar side-chain atoms and average number of atom contacts at each position for interactions between pairs of α-helices 255

164. Number of interactions between charged atoms and polar side-chain atoms and average number of atom contacts at each position for interactions between α-helices and β-sheets 256

165. Number of interactions between charged atoms and polar side-chain atoms and average number of atom contacts at each position for interactions between β-sheets and α-helices 257

LIST OF TABLES-continued

Table Page

166. Number of interactions between charged atoms and polar side-chain atoms and average number of atom contacts at each position for interactions between pairs of ß-sheets 258

167. Number of interactions between charged atoms and backbone oxygen atoms and average number of atom contacts at each position for interactions between pairs of α-helices 259

168. Number of interactions between charged atoms and backbone oxygen atoms and average number of atom contacts at each position for interactions between α-helices and ß-sheets 260

169. Number of interactions between charged atoms and backbone oxygen atoms and average number of atom contacts at each position for interactions between pairs of ß-sheets 261

170. Number of interactions between charged atoms and backbone nitrogen atoms and average number of atom contacts at each position for interactions between pairs of α-helices 262

171. Number of interactions between charged atoms and backbone nitrogen atoms and average number of atom contacts at each position for interactions between pairs of ß-sheets 263

172. Number of interactions between polar side-chain atoms and average number of atom contacts at each position for interactions between pairs of α-helices 264

173. Number of interactions between polar side-chain atoms and average number of atom contacts at each position for interactions between α-helices and ß-sheets 265

174. Number of interactions between polar side-chain atoms and average number of atom contacts at each position for interactions between ß-sheets and α-helices 266

175. Number of interactions between polar side-chain atoms and average number of atom contacts at each position for interactions between pairs of ß-sheets 267

LIST OF TABLES-continued

Table | Page

176. Number of interactions between polar side-chain atoms and backbone oxygen and average number of atom contacts at each position for interactions between pairs of α-helices 268

177. Number of interactions between polar side-chain atoms and backbone oxygen and average number of atom contacts at each position for interactions between α-helices and ß-sheets 269

178. Number of interactions between polar side-chain atoms and backbone oxygen and average number of atom contacts at each position for interactions between ß-sheets and α-helices 270

179. Number of interactions between polar side-chain atoms and backbone oxygen and average number of atom contacts at each position for interactions between pairs of α-helices 271

180. Number of interactions between polar side-chain atoms and backbone nitrogen and average number of atom contacts at each position for interactions between pairs of α-helices 272

181. Number of interactions between polar side-chain atoms and backbone nitrogen and average number of atom contacts at each position for interactions between α-helices and ß-sheets 273

182. Number of interactions between polar side-chain atoms and backbone nitrogen and average number of atom contacts at each position for interactions between ß-sheets and α-helices 274

183. Number of interactions between polar side-chain atoms and backbone nitrogen and average number of atom contacts at each position for interactions between pairs of α-helices 275

184. Number of interactions between backbone atoms and average number of atom contacts at each position for interactions between pairs of α-helices 276

185. Number of interactions between backbone atoms and average number of atom contacts at each position for interactions between α-helices and ß-sheets 277

LIST OF TABLES-continued

Table | Page

186. Number of interactions between backbone atoms and average number of atom contacts at each position for interactions between ß-sheets and α-helices 278

187. Number of interactions between backbone atoms and average number of atom contacts at each position for interactions between pairs of ß-sheets 279

188. Number of interactions between backbone atoms and hydrophobic side-chain atoms and average number of atom contacts at each position for interactions between pairs of α-helices 280

189. Number of interactions between backbone atoms and hydrophobic side-chain atoms and average number of atom contacts at each position for interactions between α-helices and ß-sheets 281

190. Number of interactions between backbone atoms and hydrophobic side-chain atoms and average number of atom contacts at each position for interactions between ß-sheets and α-helices 282

191. Number of interactions between backbone atoms and hydrophobic side-chain atoms and average number of atom contacts at each position for interactions between pairs of ß-sheets 283

192. Number of interactions between hydrophobic side-chain atoms and average number of atom contacts at each position for interactions between pairs of α-helices 284

193. Number of interactions between hydrophobic side-chain atoms and average number of atom contacts at each position for interactions between α-helices and ß-sheets 285

194. Number of interactions between hydrophobic side-chain atoms and average number of atom contacts at each position for interactions between ß-sheets and α-helices 286

195. Number of interactions between hydrophobic side-chain atoms and average number of atom contacts at each position for interactions between pairs of ß-sheets 287

LIST OF TABLES-continued

Table | Page

196. Number of occurrences of interactions between Ala ß-carbons (position -4) and residues with hydrophobic side-chains at positions relative to nearest backbone contacts. Interactions are between two α-helices with crossing angles between -150° and -180°. Average number of atom contacts are included 289

197. Number of occurrences of interactions between Ala ß-carbons (position -3) and residues with hydrophobic side-chains at positions relative to nearest backbone contacts. Interactions are between two α-helices with crossing angles between -150° and -180°. Average number of atom contacts are included 290

198. Number of occurrences of interactions between Ala ß-carbons (position 0) and residues with hydrophobic side-chains at positions relative to nearest backbone contacts. Interactions are between two α-helices with crossing angles between -150° and -180°. Average number of atom contacts are included 291

199. Number of occurrences of interactions between Ala ß-carbons (position +3) and residues with hydrophobic side-chains at positions relative to nearest backbone contacts. Interactions are between two α-helices with crossing angles between -150° and -180°. Average number of atom contacts are included 292

200. Number of occurrences of interactions between Ala ß-carbons (position +4) and residues with hydrophobic side-chains at positions relative to nearest backbone contacts. Interactions are between two α-helices with crossing angles between -150° and -180°. Average number of atom contacts are included 293

201. Number of occurrences of interactions between Ala ß-carbons (position +7) and residues with hydrophobic side-chains at positions relative to nearest backbone contacts. Interactions are between two α-helices with crossing angles between -150° and -180°. Average number of atom contacts are included 294

LIST OF TABLES-continued

Table | Page

202. Number of occurrences of interactions between Ala ß-carbons (position +11) and residues with hydrophobic side-chains at positions relative to nearest backbone contacts. Interactions are between two α-helices with crossing angles between -150° and -180°. Average number of atom contacts are included 295

203. Ratios of atom contacts/atom for each residue type for interactions between hydrophobic side-chain atoms. Ratios are for total protein contacts 297

204. Ratios of atom contacts/atom for each residue type for interactions between hydrophobic side-chain atoms. Ratios are for interactions between secondary structures 299

205. Ratios of atom contacts/atom for each residue type for interactions between oppositely charged atoms. Ratios are for total protein contacts 301

206. Ratios of atom contacts/atom for each residue type for interactions between oppositely charged atoms. Ratios are for interactions between secondary structures 303

207. Ratios of atom contacts/atom for each residue type for interactions between atoms of like charge. Ratios are for total protein contacts 305

208. Ratios of atom contacts/atom for each residue type for interactions between atoms of like charge. Ratios are for interactions between secondary structures 307

209. Ratios of atom contacts/atom for each residue type for interactions between charged atoms and polar side-chain atoms. Ratios are for total protein contacts 309

ABSTRACT

In an attempt to develop a system to aid protein structure prediction and the design of de novo proteins, residue-residue interactions in globular proteins are analyzed. Computer programs were written which studied interactions between charged atoms, polar side-chain atoms, polar backbone atoms, total backbone atoms and hydrophobic side-chain atoms. The data from these studies are analyzed and improvements in programs and new avenues of investigation are suggested.

Programs were written which would evaluate the data from the above programs. These programs have been designed to analyze the data from several perspectives. Interactions between residues were studied for the entire protein, between secondary structures in general, and between the following combinations of secondary structure pairs: 1)α-helix - α-helix, 2)α-helix - ß-sheet, 3)ß-sheet - α-helix, 4)ß-sheet - ß-sheet. For interactions between hydrophobic side chain atoms, the data were analyzed further. Interactions between pairs of residues were studied according to the angles at which the secondary structures crossed. The number of occurrences of each amino acid at each position relative to the nearest backbone contact are presented for all angle ranges. Additionally, the number of residue-residue contacts and average number of atom contacts are given for each amino acid at each position. To demonstrate the capabilities of the system to analyze data, Ala was studied for interactions between pairs of α-helices with crossing angles of -150° to -180°. For Ala at each position relative to nearest backbone atom contact, all atoms which interacted with Ala were listed according to their relative position to the nearest backbone atom contact, the number of residue-residue interactions the residue made with Ala, and the average number of atom contacts made between the two residues.

The results of the above analyses were evaluated for the heat stable protein thermolysin. The analysis of this thermophilic protein was undertaken in an attempt to provide insight into protein stability with the ultimate goal of assisting in the design of novel proteins. It is suggested that triads of α-helices, all of which interact with each other, contribute to the stability of thermolysin.

INTRODUCTION

Protein Folding

Since Anfinsen's seminal work with the renaturation of ribonuclease, it has been widely believed that the primary sequence of a protein is sufficient to determine protein structure. The prediction of protein tertiary structure remains elusive. The advent of protein X-ray crystallography has accelerated insight into factors affecting protein structure.

Energy of Folding

The folded shape of a protein is a complex interplay between enthalpy and entropy as the protein assumes a low energy conformation. In spite of daunting efforts to predict the native state of a protein from energy calculations, there has been little success. This arises in large part from the difficulty in assigning correct enthalpic and entropic contributions to the free energy of the system. Computational time required to accurately assign energy terms to all variables appears to be prohibitive. This has necessitated that alternative avenues of investigating protein structure and stability be investigated. A major effort in this area is being assigned to pattern recognition, under the premise that rather than attempting the calculation of energy ab initio, an understanding of favorable

interactions may be more accurately ascertained through investigation of known protein structures.

The folding of a protein is widely believed to be a thermodynamic process involving the interplay between the aqueous solvent and regions of proteins which vary in hydrophobicity. While the effects of water on protein structure are not studied explicitly in this work, it is recognized that water is central in the determination of protein structure.

When hydrophobic groups are introduced into water, enthalpic and entropic changes occur. The creation of a hole in the hydrogen-bonded water structure is enthalpically unfavorable(3). However, there are favorable enthalpic contributions due to London forces between the non-polar group and water. It has been determined that the attractive force between nonpolar molecules and water is comparable to that between non-polar molecules(3). The result is that the net enthalpy of interaction of a hydrophobic group with water may be favorable.

When nonpolar groups are introduced into water, a cage of hydrogen bonded water molecules is formed. This causes the entropy of the system to decrease. Hydrophobic groups tend to self-associate in water to decrease their total solvent accessible surface area and minimize the unfavorable decrease in entropy(3). This is the basis of the dogma that hydrophobic side chains bury in the protein interior while charged and polar side-chains tend to reside on the protein surface. The amount of buried area has been related to stability(12) and the amount of solvent exposed surface can be predicted(24). Through computer

simulations, it has been shown that helices and sheets form upon polymer compaction(9) and the amount of secondary structure increases as the chain becomes increasingly compact(44).

Water affects conformations of side chains on the protein surface and is often found in active sites, where it is usually involved in mediating electrostatic strain in enzymatic activity. Water is often ordered in X-ray protein structures and is often found to interact with secondary structures where they protrude from the hydrophobic interior(2). The groups interacting with bound water tend to be charged and the affect of water at the ends of secondary structures is likely to affect the conformations of turns.

Charged Groups

The potentially charged groups in proteins are the N- and C-termini, and side chain atoms from Asp, Glu, Cys, Tyr, Lys, Arg, and His. Of these side chains, only Asp, Glu, Lys and Arg are normally charged. At physiological pH, less than one in ten His residues will carry a positive charge. This is supported by a study of potential salt-bridges where it was determined that salt-bridges involving His marginally contributed to stability(60).

The strength of interactions between ion pairs decreases with 1/r, where r is the distance separating the centers of the atoms in question. The dependence of the dielectric constant upon groups surrounding charged atoms makes accurate prediction of the strength of charged interactions difficult. This difficulty is exacerbated when shielding effects of salt groups are taken into account. These

factors adversely affect the accuracy of models which extrapolate the cost of burying charged groups from water to non-polar solvents to the cost of burying amino acids in folded proteins. Perhaps more telling is the topological distribution of charged groups on the protein surface. An average of only one pair of interacting charged groups per 150 residues are found to be buried in the hydrophobic interior(28). When charged groups are found in the protein interior, they are usually located within 4 Å of an oppositely charged group and if not, they are nearly always hydrogen bonded to a polar group(20).

Much of the insight into contributions of charged groups on protein stability has been garnered through site-directed mutagenesis experiments and analysis of X-ray structures. Currently, charged interactions are not considered to be dominant in the determination of the folded state(9). From a statistical point of view, this is exemplified by the poor evolutionary conservation of charged groups, with the noted exception of charged groups located in binding sites and active sites(28). However, ion pairs in thermophilic proteins are often the predominant difference between these proteins and their mesophilic counterparts(45,62). There is near consensus agreement that charged groups do affect protein stability.

The position of charged groups seems to be important in determining whether they contribute to stabilization or destabilization. Intra-helical salt bridges are often formed between side chains of oppositely charged residues separated by either three or four residues(60). This was demonstrated when two peptides were synthesized, one composed of repeating blocks of glu_4 and lys_4 and the other

composed of repeating blocks of glu_2 and lys_2(41). The peptide which repeated after four residues was stable while the peptide which repeated after two residues was not stable. The authors concluded that ion pairing was responsible for the differences in stability. While these interactions are considered to stabilize helix formation, the magnitude of the contribution to protein stability is subject to debate. The insertion of glutamate and lysine residues four residues apart on a surface helix of barnase has been estimated to contribute only 0.2 kcal/mol to stability(38). This contribution is at the lower end of salt-bridge stabilization, where estimates usually range from three to five kcal/mol. Addition of charges which cannot be well solvated and do not form intramolecular salt-bridges result in the destabilization of proteins(54,55). Clustering of positively charged groups in both bovine pancreatic trypsin inhibitor and phage T4 lysozyme result in destabilization(14,53,57). The distance over which charged groups mediate their effect is large. In subtilisin, charged groups up to 20 Å from a His residue affect the pK of that His residue(4).

Charged groups are often found at the ends of helices. There is a tendency for negatively charged groups to be found at the amino-termini of helices and for positively charged groups to be found at the carboxy-termini. This may be a result of the evolutionary drive to stabilize the helical dipole, which carries +0.25 e^- at the amino end and -0.25 e^- at the carboxyl end. As with nearly all other calculations of stability, there is dispute as to the magnitude of stabilization. It has been suggested, from work done on the four-helix bundle protein myohemerythrin, that interactions between charged groups and the helical dipole have been

overestimated, at least with respect to the stability of myohemerythrin(56). However, in this study, the effect of charged interactions between groups on different helices was not taken into account. Likewise, the stable de novo protein designed by DeGrado and co-workers, based upon the four-helix bundle proteins myohemerythrin, cytochrome c′, and tobacco mosaic coat protein(43), analyzed the placement of charged groups solely from the standpoint of stabilization of the helical dipole. However, the antiparallel sense of helices in these models places the positively charged amino acids at the C-terminus of one helix in close proximity with negatively charged amino acids at the N-terminus of an adjacent helix. There is a paucity of information regarding the study of charged groups between the ends of different secondary structures.

A study of the preponderance of hydrogen bonds in protein structures(59) found that 80% of Asp side-chains and 72% of Glu side-chains formed hydrogen bonds through both carboxylate oxygens. For Arg side-chains, 33% had all three guanido nitrogen groups involved in hydrogen bonds, 35% had two groups hydrogen bonded, and 25% had only one group involved in hydrogen bonds. Of the hydrogen bonded guanido nitrogens, there was a preference for the terminal NH_2 groups to be involved rather than the ε-NH group. For Lys, 24% formed no hydrogen bonds, 36% formed one hydrogen bond, 19% formed two hydrogen bonds, 14% formed three hydrogen bonds, and 7% formed more than three hydrogen bonds, presumably through shared hydrogen bonds. Glu and especially Asp were characterized by making a high number of contacts with backbone

nitrogen atoms. Interestingly, Lys residing on the surface made fewer contacts with water than all of the other charged groups as well as Thr, Asn, Gln, and His.

The Hydrophobic Interior

The interior of a protein is dominated by hydrophobic amino acids. These amino acids utilize their wide range of dihedral angles to form low energy(7,25,32,57) close-packing structures. While some regions pack more closely than others, the overall packing density for proteins is comparable to the packing density of small organic molecules(5,22). There are very few empty regions the size of a methyl group within the protein interior(8,17).

The importance of hydrophobic packing is evidenced by the high evolutionary conservation of hydrophobic residues(58). There does not appear to exist a preference for the degree of burial of different hydrophobic groups. Large hydrophobic groups were not found to be buried more frequently or more fully than small hydrophobic groups(18). Specifics of hydrophobic interactions will be addressed more fully in the Introduction section dealing with packing interactions.

In spite of the daunting task of creating alternate hydrophobic regions in the protein interior, a number of attempts have yielded impressive results. The hydrophobic core of T4 lysozyme was altered without deleterious results by rational design. Several multiple mutants were made which destabilized the protein by less than 0.2 kcal/mol(42). Shuffling of residues in the hydrophobic core of a DNA binding protein resulted in mutants which retained partial activity(49). This experiment and another which modified λ repressor(36) suggest that the critical

determinant in these mutational studies is that the hydrophobic core remain hydrophobic.

Hydrogen Bonds

Not surprisingly, the contribution of hydrogen bonds upon protein stability has been widely debated. Hydrogen bonds have often been assumed to contribute little to protein stability. This was based upon the assumptions that there would be equal numbers of hydrogen bonds formed in the folded and unfolded states and that hydrogen bonds are energetically equivalent, with hydrogen bonds contributing primarily to specificity of folding. This theory has been questioned on the grounds that, in a protein, the ability to make and break hydrogen bonds is greatly reduced, resulting in an effective increase in the concentrations of hydrogen bonds(4).

The importance of hydrogen bonds can perhaps most accurately be assessed by two complementary approaches: The first is to analyze hydrogen bonds in folded proteins and determine what effect they might have in stabilizing the folded state of proteins. The second approach is site directed mutagenesis where hydrogen bonds are eliminated or hydrogen bonds are added where they did not previously exist. Site directed mutagenesis is often limited by steric constraints which often result in destabilization of the folded state.

An extensive study on the role of hydrogen bonding was undertaken on phage T4 lysozyme. Thr 157 was found to be important in the stabilization of T4 lysozyme. Substitution of Thr with Ile reduced the stability of lysozyme by approx

3 kcal/mol(29,31). Clearly, the size difference between Ile and Thr could not be ruled out as the dominating effect. However, Thr formed a hydrogen bond with the main-chain amide of Asp 159. Twelve additional mutants were created with substitutions at position 157(47). The five most stable mutagens all formed hydrogen bonds with the main chain amide nitrogen at Asp 159.

This study also showed that water can stabilize structures via ordered hydrogen bond formation. Gly 157 bound a water molecule in the crevice which resulted from its replacement of Thr and was nearly as stable as the native protein.

In another study involving phage T4 lysozyme, Pro 86 was substituted with 10 individual amino acids(4). Of these substitutions, Ser and Cys were shown crystallographically to form hydrogen bonds with the side chain of Gln 122. However, these mutants were no more stable than the mutants which did not form hydrogen bonds, suggesting that hydrogen bonds are not a panacea for protein stabilization. This raises a question central to the premise of this work; are some contributions to stability more important than others?

Secondary Structures

The two predominant secondary structures, α-helices and ß-sheets are characterized by their main chain dihedral angles, Φ and Ψ. For α-helices, the average Φ and Ψ angles are approximately -60° and -50° respectively, which places them in the lower left-hand corner of a Ramachandran plot. For ß-sheets,

the range of Φ-Ψ angles is much broader, and they fall into the upper left-hand quadrant of a Ramachandran plot.

α-Helices

The α-helix was first predicted in 1951 by Pauling(50). Using small-molecule crystal structures, he predicted the α-helix to be stable based upon geometrical factors. α-helices are characterized by 3.6 residues/turn with hydrogen bonds formed between C=O of residue n and the N-H group of residue $n + 4$. The rise per residue along the helix axis is approximately 1.5 Å. α-helices nearly always have a right-handed twist and the average length of an α-helix is 10 amino acids. Amphiphilicity, which is so often observed for α-helices, is a consequence of their preference for the protein surface, resulting in one side being embedded into the hydrophobic protein interior and the other exposed to aqueous solvent. The repeated hydrogen-bonds in α-helices orient all of the helical dipoles in the same direction resulting in a helical net dipole.

α-helices appear to be autonomously stable building blocks. This is supported by the crystal structure of melittin(10) and experiments which demonstrate the stability of isolated protein helices in water(46). As will be discussed, α-helices appear to play a major role in the stabilization of the heat stable thermolysin.

ß-Sheets

ß-sheets, by definition, are comprised of more than one ß-strand. The orientation of the strands in a ß-sheet can be parallel, antiparallel, or mixed, with bias against mixed ß-sheets. The average length for a ß-strand is 5 to 10 amino acids, with longer strands being more common in antiparallel sheets rather than in parallel sheets(11). Adjacent α-carbons alternate sides of the backbone and α-carbons two residues distant are separated by approximately 7 Å(51). ß-sheets are characterized by a relatively flat surface on which to pack. There is a preference for side chains from groups adjacent to each other on neighboring strands to pair charged groups of opposite charge, to pair hydrophobic groups together, and to pair side chains branched at ß-carbons with side chains which are unbranched in the ß-carbon position(2).

Hydrogen bonding in ß-sheets is much more directly aligned in antiparallel ß-sheets than parallel ß-sheets, which most likely contributes to the higher stability of antiparallel ß-sheets. The nature of alignment of antiparallel ß-sheets orients two adjacent, direct hydrogen bonds close to each other with a wider gap between these and the next set of hydrogen bonds. Parallel ß-sheets are characterized by a more even distribution of diagonal hydrogen bonds. Both parallel and antiparallel sheets share the common trait of having an overall right-handed twist(21).

Parallel ß-sheets almost always are comprised of at least five strands and these strands are generally removed from solvent on both sides(2). This is

contrasted to anti-parallel ß-sheets which often consist of just two strands in which one side of the sheet is exposed to solvent. Both ß-sheets display a tendency to be more hydrophobic in the central regions relative to end regions(23), resulting in central regions being embedded in the hydrophobic interior of the protein.

Motifs

Secondary structures combine to form motifs, regular identifiable patterns of protein structures. These motifs are influenced by a number of factors, such as constraints on their surface area and how different secondary structures pack. The packing is dominated by side-chains, which strive to close-pack. Three basic types of packing structures constitute the majority of packing observed in structural motifs. The combinations which are most common are helix-turn-helix, beta-turn-beta, and beta-alpha-beta structures. As their names imply, these motifs compose most of the methods of combining α-helices and ß-sheets. While helix-turn-helix motifs clearly connect two α-helices, beta-turn-beta motifs commonly join two ß-strands together in an antiparallel orientation, and beta-alpha-beta motifs are the building blocks of α/ß open-twisted sheets and α/ß barrels.

Beta-Alpha-Beta Motifs

In beta-alpha-beta structures, the α-helix packs nearly parallel to the ß strands(26). The ß-strands in α/ß proteins average seven amino acids in length and α-helices average 14 residues in length. ß-bulges are seldom found in these proteins. α-helices involved in beta-alpha-beta structures tend to be amphiphilic,