Exploring Secondary Structure Elements

5.1 Background – 87

5.2 Locating Secondary Structure in Papain and RNase A – 88

5.3 Consequences of Secondary Structure Formation – 92

5.4 Assigning Secondary Structure – 92

5.4.1 Secondary Structure Assignment by PyMOL – 97

5.5 Exploring β-Sheets – 99

5.5.1 Looking into the β-Strand Assigned Only by DSSP – 102

5.5.2 Investigating Parallel β-Strands – 102

5.6 Turns – 104

5.6.1 β-Turns – 104

5.6.2 γ-Turns – 107

5.6.3 The Helical 3_{10} Turn (or a Type III β-Turn) – 108

5.6.4 The α-Turn with Bonding from Residue *i* to *i* + 4 – 108

5.6.5 The Complexity of Turns in Proteins – 109

5.7 Exploring Helical Structures of Proteins – 109

5.7.1 α-Helices – 109

5.7.2 Variations Found at the N- and C-Termini of α-Helices – 112

5.7.3 Arrangements of the Side-Chains in α-Helices – 114

5.7.4 3_{10} Helices – 114

5.7.5 Pi Helices – 115

5.8 Conformational Preferences of Amino Acids – 115

5.9 Summary of PDB Information Discussed in This Chapter – 118

T. Skern, *Exploring Protein Structure: Principles and Practice*, Learning Materials in Biosciences,
https://doi.org/10.1007/978-3-319-76858-8_5

5.10 Summary of PyMOL Commands Introduced in This Chapter – 118

5.10.1 General and Settings – 118
5.10.2 Organization – 118
5.10.3 Viewing – 118
5.10.4 Selecting – 118
5.10.5 Coloring – 119
5.10.6 Changing – 119

5.11 Further Reading – 119

5.11.1 Books – 119
5.11.2 Online Resources – 119

5.12 Exercises – 120

References – 121

Wool gave me a glimpse of the loom on which the web of life was woven. William Astbury

What You Will Learn in This Chapter

In the previous chapter, we investigated the structure of the polypeptide backbone in the two main types of secondary structure elements, the α-helix and the β-sheet. In this chapter, the secondary structure elements themselves will be presented in more detail. First, we will learn how residues are assigned to such elements by different algorithms. This will reveal that the assignments of secondary structures are subject to some uncertainty and may vary depending on whether an algorithm uses torsional angles or hydrogen-bonding possibilities or both to make assignments. We will compare and contrast the hydrogen-bonding properties of α-helices and β-sheets and examine some variations of these structures that occur in protein structures. The chapter will also introduce the various types of turns that link the secondary structure elements together and show how the turns can be characterized by their torsion angles and hydrogen-bonding properties. A discussion of preferences of the amino acids to be found in α-helices, β-sheets, or turns will round off the chapter (◘ Table 5.1).

5.1 Background

Suggestions on the arrangement of amino acids in proteins became concrete between the 1930s and the early part of the 1950s. One of the leaders in the field was William Astbury, who worked on the structure of wool proteins at the University of Leeds. Leeds and the surrounding area were the center of the British wool industry. Astbury observed distinct changes in the fiber diffraction patterns of wool keratin when it was stretched. He termed the non-stretched form the α-form of keratin and considered it helical. In contrast, the stretched form was fully extended; he termed this the β-form of keratin (Astbury and Woods 1930). In the 1930s, Astbury proposed the importance of main-chain hydrogen bonds in determining the structure of proteins (Astbury and Woods 1934). A further advance was made by Linus Pauling in 1940 in a paper on the structure and synthesis of immunoglobulin molecules. The paper contained a drawing of an anti-parallel β-sheet with hydrogen bonds between the carbonyl oxygen and amide nitrogen atoms (Pauling 1940); subsequently, Pauling also proposed the existence of the parallel β-sheet (Pauling and Corey 1951).

◘ **Table 5.1** Structures and their PDB identifiers examined in this chapter

PDB entry identifier	PDB entry title	Organism	Reference
9pap	Structure of papain refined at 1.65 Angstroms resolution	*Carica papaya*	Kamphuis et al. (1984)
7rsa	Structure of phosphate-free ribonuclease A at 1.26 Angstroms	*Bos taurus*	Wlodawer et al. (1988)
1q21	Crystal structures at 2.2 Angstroms' resolution of the catalytic domains of normal ras protein and an oncogenic mutant complexed with GSP	*Homo sapiens*	Tong et al. (1991)
1din	Dienelactone hydrolase at 2.8 Angstroms	*Pseudomonas knackmussii*	Pathak and Ollis (1990)

Drawing plausible structures for an α-helix turned out to be more difficult. In 1943, Maurice Huggins, an American chemist, published a long and detailed theoretical study of putative structures for α-helices and β-sheets (Huggins 1943). Huggins' description of the peptide bond gives an idea of the state of the field at this time. Nevertheless, the credit for the determination of the structure of the α-helix went to Linus Pauling and colleagues. Pauling came to the solution through model building which he started while he was recovering in bed from a cold in Oxford (Hager 1998). To pass the time, he began to draw paper models of polypeptide chains to examine ways of arranging the main-chain atoms. As a basis for his models, Pauling used data from X-ray studies of crystallized amino acids and dipeptides. The distance between the carbon and nitrogen atoms in the dipeptides suggested that the peptide bond possessed a partial double bond character and a planar structure. Pauling's knowledge of resonance theory of the chemical bond also implied a partial double bond character for the peptide bond. Importantly, therefore, the models were drawn with a planar peptide bond and thus formed the basis for two helical structures that were published in April 1951 (Pauling et al. 1951); one was the correct structure of the α-helix. Pauling was obviously very active at this time in examining and modeling the structures of polypeptides; in May 1951 alone, he published six further papers back to back to follow the one cited above.

There was however no experimental evidence whether the helix proposed by Pauling was actually present in a globular protein. Fortunately, soon after the paper's publication, such support was forthcoming. Max Perutz read Pauling's seminal paper in PNAS (Pauling et al. 1951) one Saturday morning. He realized that he would only be able to observe X-ray reflections from such a helix if he rotated his X-ray apparatus 90° from its standard position. The very same afternoon, he took a horse hair from his desk drawer, placed it at right angles to the X-ray beam, and collected the reflections on film. After developing the film, he was able to see a characteristic diffraction pattern at 1.5 Å, which had been predicted by the Pauling model. Perutz published the results in a brief report in Nature on June 30, 1951 (Perutz 1951), just 3 months after Pauling's paper was published.

Why did Perutz and his colleagues in Cambridge not find the solution to the structure of the α-helix? It seems that Perutz was not aware of or did not consider the planar nature of the peptide bond which was a vital part of Pauling's models. The situation was reversed in the race for the structure of the DNA helix when the Cambridge group, in contrast to Pauling, had been informed of the correct resonance structures of the DNA bases. The structures of the α-helix and the DNA helix formed the very foundations of molecular biology. The knowledge of the correct chemistry of the respective building blocks was thus a prerequisite for both discoveries.

5.2 Locating Secondary Structure in Papain and RNase A

We mentioned the secondary structure elements in ▶ Chaps. 3 and 4 without looking at them in detail. Let us again draw the structures of papain and RNase A, this time coloring the molecules according to the secondary structure elements (◘ Fig. 5.1a, b). Using a simple PyMOL command ("find polar contacts"), we can then examine the main-chain interactions that maintain them (◘ Figs. 5.1b and 5.2b). ▶ Box 5.1 shows you how to make the two images in ◘ Fig. 5.1 in the same PyMOL session so that they can be saved in the same orientation. These instructions can of course be used for RNase A, except that the PDB code will be 7rsa.

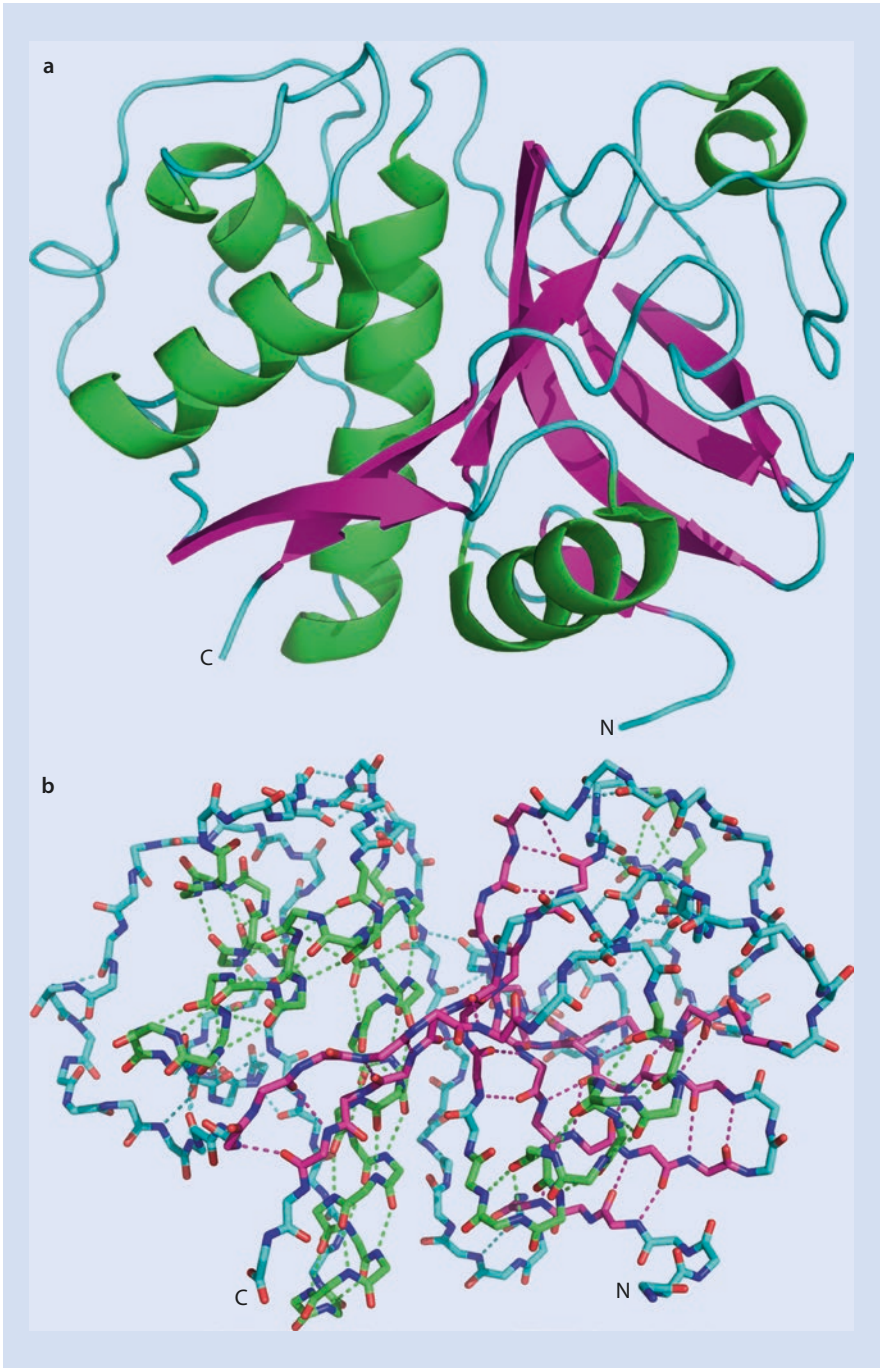


Fig. 5.1 The secondary structure elements of papain and the main-chain interactions that maintain them. **a** α-helices are in green, β-strands are in magenta, and coils (or loops) connecting the elements are in cyan. **b** Stick representation of the main chain of papain. Carbon atoms are green in α-helices, magenta in β-strands, and cyan in coils (or loops). Oxygen atoms are red, nitrogen blue. Polar contacts between main-chain atoms are color-coded according to their secondary structure assignment

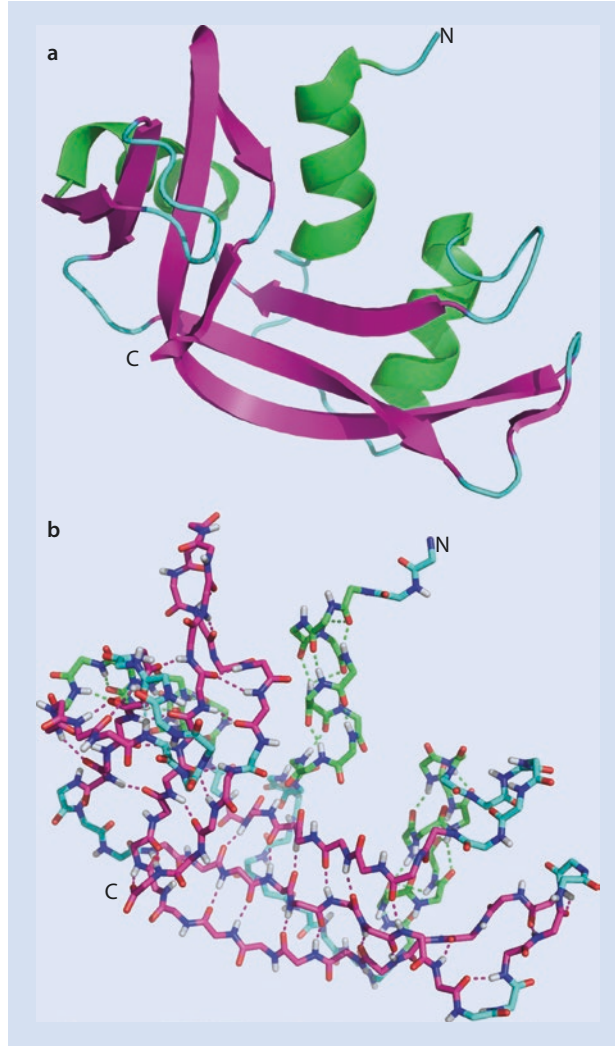


Fig. 5.2 The secondary structure elements of RNase A and the main-chain interactions that maintain them. **a** α-helices are in green, β-strands are in magenta, and coils (or loops) connecting the elements are in cyan. **b** Stick representation of the main chain of RNase A. Carbon atoms are green in α-helices, magenta in β-strands, and cyan in coils (or loops). Oxygen atoms are red, nitrogen blue, and hydrogen atoms white. Polar contacts between main-chain atoms are color-coded according to their secondary structure affiliation

Figures 5.1 and 5.2 contain a great deal of information and require close attention. We will look specifically at the interactions found in the helices, strands, and loops later in the chapter. First, though, let us think about why we asked PyMOL to find polar contacts and not hydrogen bonds. To find hydrogen bonds, PyMOL would need information about the charges on ionizable groups (e.g., what are the ionization states of the histidine residues?). Further, if the hydrogen atoms are not present in the structure, then PyMOL has to suggest suitable coordinates for them that may not be accurate. In other words, asking PyMOL to automate the search for hydrogen bonds is not sensible. Instead, it is much more reliable (and instructive) to let PyMOL propose polar contacts and then decide yourself the exact nature of the interaction between the two atoms (i.e., is there a hydrogen bond or an electrostatic interaction, etc.). To see some examples of polar contacts that are not hydrogen bonds, zoom into residues Val32, Thr33, and Ile 34 (zoom i. 32–34). You can see two amide nitrogens making two polar contacts. As each amide nitrogen can only

Box 5.1 Drawing the Secondary Structure Elements and Main-Chain Polar Contacts as in ◘ Figs. 5.1 and 5.2

```
#         bring up the entry into PyMOL
Open the PyMOL program
PyMOL>    fetch 9pap (or 7rsa)
PyMOL>    as cartoon
#         color according to secondary structure (ss) (see ► Sect. 3.10.5)
PyMOL>    color green, ss h
PyMOL>    color magenta, ss s
PyMOL>    color cyan, ss l+
#         show the sequence of the entry colored according to secondary structure
PyMOL>    set seq_view, on
Mouse>    for 9pap, rotate the molecule so that the domain without
          α-helices is on the right and the N- and C-termini are at
          the bottom
>         File, Save Session as "papain_5_1a", Quit
Mouse>    for 7rsa, rotate the molecule so that the N-terminus is on
          the top right and the C-terminus is on the left and point-
          ing to the back
#         these are the drawings in part A. If you wish to print, then set the background
          to white now. Save the file under a different name to prevent overwriting it by
          mistake
>         File, Save Session as "papain_5_1b"
#         to make the drawings in part B, use the mouse to hide the ribbon by clicking on
          9pap or 7rsa on the GUI on the right
#         show the main chain on top of the cartoon by defining a new object (for 7rsa,
          modify file names accordingly)
PyMOL>    create pap_mc_sticks, 9pap
PyMOL>    hide everything, pap_mc_sticks
#         show and color the main chain atoms according to element
PyMOL>    show sticks, pap_mc_sticks and name n+ca+c+o and not hetatm
PyMOL>    select pap_mc_sticks and ss h
PyMOL>    util.cbag sele
GUI       sele, actions, find, polar contacts, just intra_main chain
GUI       sele_polar_conts, actions, rename, helix_pc
#         color the dashes of the polar contacts for helices, care to use "by rep"
GUI       helix_pc, color, by rep, dashes, greens, green
PyMOL>    select pap_mc_sticks and ss s
PyMOL>    util.cbam sele
GUI       sele, actions, find, polar contacts, just intra_main chain
GUI       sele_polar_conts, actions, rename, strand_pc
#         color the dashes of the polar contacts for strands, care to use "by rep"
GUI       strand_pc, color, by rep, dashes, magentas, magenta
PyMOL>    select pap_mc_sticks and ss l+
PyMOL>    util.cbac sele
GUI       sele, actions, find, polar contacts, just intra_main chain
GUI       sele_polar_conts, actions, rename, loop_pc
#         color the dashes of the polar contacts for loops, care to use "by rep"
GUI       loop_pc, color, by rep, dashes, cyans, cyan
#         change the background color
PyMOL>    bg white
#         save the session
> File, Save Session "papain_5_1b", Quit
```

donate one hydrogen bond, only one polar contact (the straight one in the direction of the helix) is a hydrogen bond.

In summary, it is important to understand what defines a hydrogen bond and to understand the basis of any automation protocol. Further, these difficulties in assigning hydrogen bonds from the data in PDB files illustrate why the mutagenesis experiments of Fersht et al. (1985) (mentioned in ▶ Sect. 3.5) were so groundbreaking. For the first time, they showed that the removal of the potential to form a single hydrogen bond really did affect the activity of an enzyme.

5.3 Consequences of Secondary Structure Formation

Figures 5.1b and 5.2b show the large number of hydrogen bonds that maintain the structure of a protein. In addition to the maintenance of the structure, the hydrogen bonds are responsible for making the structure possible in the first place as they neutralize the charges on peptide bond. Without this charge neutralization, the polypeptide chain would be unable to be part of a hydrophobic core. Indeed, the formation of these intra-main-chain hydrogen bonds plays an important role in the folding of a protein. If a protein is completely unfolded, all of the amide hydrogens and carbonyl nitrogens of the polypeptide chain would be hydrogen bonded to water molecules. Looked at in terms of energy (i.e., enthalpy), there are fewer hydrogen bonds in a folded protein than in an unfolded one because not all peptide bonds are hydrogen bonded. This would seem to be a negative influence on protein folding and stability. However, some strong hydrogen bonds may compensate to some extent for this effect; further, evidence from site-directed mutagenesis of proteins also indicates a role for hydrogen bond formation in folding and stabilizing proteins.

Looked at in terms of disorder or entropy, the formation of hydrogen bonds is however advantageous. The water molecules hydrogen bonded in the unfolded state are fixed and hence have a low entropy value; their release through the formation of hydrogen bonds in the secondary structure frees the water molecules and increases the entropy of the system. The loss of entropy of the protein through its folding is much less than compared to the entropy of the freed water molecules.

Thus, the formation of intra-chain hydrogen bonds is crucial to protein folding and stability even though the exact contributions of enthalpy and entropy remain unclear. You can read more on this subject in a very comprehensive review (Pace et al. 1996) and in a more concise description that is available on the internet (▶ http://faculty.smu.edu/svik/6312/Lectures/8Feb.html).

5.4 Assigning Secondary Structure

The PyMOL drawings in Figs. 5.1 and 5.2 show the secondary structure elements of papain. But how does PyMOL know when to draw a helix or a strand or a turn? Where does PyMOL obtain its information and how can we be sure that it is an accurate description? The first place to look for an accurate description of the secondary struc-

a

	A		B			C	D	E
HELIX	1	L1	SER A 24	GLY A 43		1	DISRUPTION IN THE CENTER	20
HELIX	2	L2	GLU A 50	ASP A 57		1		8
HELIX	3	L3	TYR A 67	TYR A 78		1		12
HELIX	4	R1	ASN A 117	GLN A 128		1		12
HELIX	5	R2	GLY A 138	LEU A 143		1		6

b

	F	G	H			I	J					
SHEET	1	S1	4	VAL A 5	TRP A 7	0						
SHEET	2	S1	4	VAL A 164	GLY A 167	-1	O	TYR A 166	N	VAL A 5		
SHEET	3	S1	4	TYR A 170	ASN A 175	-1	O	LEU A 172	N	VAL A 164		
SHEET	4	S1	4	GLY A 185	ARG A 191	-1	O	ILE A 187	N	ILE A 173		
SHEET	1	S2	4	ASP A 108	VAL A 113	0						
SHEET	2	S2	4	SER A 206	VAL A 210	-1	O	TYR A 208	N	ARG A 111		
SHEET	3	S2	4	VAL A 130	LEU A 134	-1	O	SER A 131	N	PHE A 207		
SHEET	4	S2	4	ASP A 158	ALA A 163	-1	O	VAL A 161	N	VAL A 132		

Fig. 5.3 Helix **a** and sheet **b** records from the PDB entry for 9pap. Key: **A.** Number and identifier of helix. **B.** Residue names, chains, and numbers of first and last residue in the helix. **C.** Type of helix (see Table. 5.2). **D.** Comment (see text) **E.** Number of residues in the helix. **F.** Number and identifier of β-sheet. **G.** Number of strands in β-sheet **H.** Residue names, chains, and numbers of first and last residues in the β-strand. **I.** Direction of β-strand relative to previous one, 0 for the first strand, −1 for anti-parallel, 1 for parallel. **J.** Atoms of the β-strands that form a hydrogen bond

Table 5.2 Types of helix referred to in the PDB. You can also find this information in a useful introduction to the PDB at the UCSF web page: ▶ https://www.cgl.ucsf.edu/chimera/docs/UsersGuide/tutorials/pdbintro.html

Helix types			
1	Right-handed alpha (default)	6	Left-handed alpha
2	Right-handed omega	7	Left-handed omega
3	Right-handed pi	8	Left-handed gamma
4	Right-handed gamma	9	2/7 ribbon/helix
5	Right-handed 3/10	10	Polyproline

ture assignment is the PDB file. Call up the PDB header for 9pap from the PDB and search for "helix." You will find a table describing the protein's secondary structure as shown in Fig. 5.3.

The information for the record "HELIX" is relatively straightforward and intuitive, except for that in column C. This integer can vary from 1 to 10 and describes the type of helix as illustrated in Table 5.2. The comment indicates that the turn in the middle of the helix is tighter (i.e., the pitch is shorter) than at the ends of the helix.

a

	A		B			C	D	E
HELIX	1	A1	GLY A 15	ASN A 26		1		12
HELIX	2	A2	ASP A 69	GLY A 75		1		7
HELIX	3	A3	THR A 87	LYS A 104		1		18
HELIX	4	A4	GLU A 126	TYR A 137		1		12
HELIX	5	A5	GLY A 151	LEU A 171		1		21

b

	F	G	H		I	J				
SHEET	1	S1	6	ASP A 38	ILE A 46	0				
SHEET	2	S1	6	GLU A 49	ASP A 57	-1	N	LYS A 42	O	LEU A 53
SHEET	3	S1	6	THR A 2	VAL A 9	1	O	LEU A 6	N	LEU A 56
SHEET	4	S1	6	GLY A 77	ILE A 84	1	O	LEU A 79	N	VAL A 9
SHEET	5	S1	6	PRO A 110	LYS A 117	1	O	VAL A 112	N	CYS A 80
SHEET	6	S1	6	PRO A 140	THR A 144	1	O	ILE A 142	N	GLY A 115

Fig. 5.4 Helix **a** and sheet **b** records from the PDB entry for p21^{ras} (1q21). The columns are explained in the legend to Fig. 5.3.

It is evident that there are ten different types of helices found in proteins. We will look at right-handed α-helices, right-handed π-helices, and right-handed 3_{10} helices later in the chapter in the section on helices.

Two columns, I and J, in the records "SHEET" in Fig. 5.3 also require some explanation. Column I indicates the direction of each strand relative to the previous one. The value is thus 0 for the first strand. A value of −1 means that the strand is anti-parallel; a value of 1 indicates a parallel one. The cell signaling protein p21^{ras} that was mentioned in Exercise 4.9.5 is a protein that has a characteristic mixed β-sheet. Compare the sheet records for this protein below (Fig. 5.4) with those for papain in Fig. 5.3. It is immediately obvious from column I that the β-sheet is a mixed one, comprising one anti-parallel and five parallel β-strands.

To understand column J in the records "SHEET," let us move on to Fig. 5.4. In the line describing the β-strand starting at Glu A 49, we note at the right "`N LYS A 42 O LEU A 53.`" This indicates that the amide nitrogen of Lys42 from strand 1 hydrogen bonds with the oxygen atom of Leu53 in strand 2. Similarly, in the line describing the strand starting at Thr A 2, we note "`O LEU A 6 N LEU A 56.`" This indicates that Leu56 of strand 2 hydrogen bonds with Leu6 of strand 3. In other words, the information indicates which amino acid of a particular β-strand hydrogen bonds with the previous one. Logically, there is no information for the first strand of each sheet. To help your understanding, Exercise 5.12.3 asks you to use the information in Figs. 5.3 and 5.4.

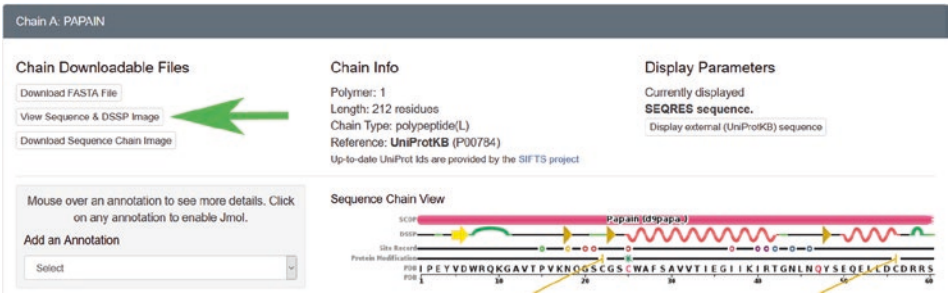


Fig. 5.5 Section of the page under "sequence tab" of the PDB entry 9pap. The green arrow marks the "View Sequence & DSSP Image" button

Where does the information on secondary structure assignments in the PDB come from? The information is either provided by the depositor or it is generated by an algorithm called DSSP ("define secondary structure of proteins") (Martin et al. 2005). DSSP is one of the oldest bioinformatics programs that is still in use. It was developed by Kabsch and Sander in 1983 (Kabsch and Sander 1983a) and employs electrostatic and hydrogen-bonding properties to define secondary structure. Over time, it has become somewhat of a gold standard for the assignment of secondary structures. However, there are many algorithms that can assign secondary structure using the same parameters as DSSP; some algorithms also include the values of the ϕ and ψ angles. Many of these algorithms were compared by Martin et al. (2005). Kabsch and Sander (1983b) also discussed the issue of accuracy in secondary structure prediction and assignment.

In addition to the information in the PDB file on assignments, each PDB entry also has a link to the DSSP output itself. You can find the link under the "Sequence" tab. Scroll down the page under the "Sequence" tab for 9pap to see the colored sequence with a host of features, including a pictorial view of DSSP under the section "Sequence chain view" (see Fig. 5.5 for a section of this page). We are interested in just the DSSP output, so click on "View Sequence & DSSP Image" (to the left of the page, marked with a green arrow in Fig. 5.5).

You will then see the sequence and DSSP output for papain (Fig. 5.6a); that for $p21^{ras}$ as an example of a protein with a mixed β-sheet is shown in Fig. 5.6b. The output from DSSP contains more than just an assignment of the β-strands and the α-helices; it includes turns, bends, and β-bridges. We will look at these later in the ▶ Sect. 5.6 of this chapter. First, let us compare the secondary structure assignments found in the PDB, DSSP, and PyMOL outputs. These are shown in Tables 5.3 and 5.4.

- H = α-helix
- B = residue in isolated β-bridge
- E = extended strand, participates in β ladder
- G = 3-helix (310 helix)
- I = 5 helix (π-helix)
- T = hydrogen bonded turn
- S = bend

a Sequence and secondary structure for 9PAP chain A

```
1    IPEYVDWRQK GAVTPVKNQG SCGSCWAFSA VVTIEGIIKI RTGNLNQYSE
       S EETTTT T      B   SSB HHHHHH HHHHHHHHHH HHS    B H

51   QELLDCDRRS YGCNGGYPWS ALQLVAQYGI HYRNTYPYEG VQRYCRSREK
     HHHHHH TTS  TTB   HHH HHHHHHHT B  BTTTS   S S     SGGG

101  GPYAAKTDGV RQVQPYNQGA LLYSIANQPV SVVLQAAGKD FQLYRGGIFV
      S SB  SEE EE  SS HHH HHHHHHHS E EEEE   SHH HHT  SSEE

151  GPCGNKVDHA VAAVGYGPNY ILIKNSWGTG WGENGYIRIK RGTGNSYGVC
         S   EE EEEEEEETTE EEEE SB TT STBTTEEEEE   SS SS GG

201  GLYTSSFYPV KN
     GTTS  EEEE
```

b Sequence and secondary structure for 1Q21 chain A

```
1    MTEYKLVVVG AGGVGKSALT IQLIQNHFVD EYDPTIEDSY RKQVVIDGET
      EEEEEEEE   TTSSHHHHH HHHHHSS  S S  TT  EEE EEEEEETTEE

51   CLLDILDTAG QEEYSAMRDQ YMRTGEGFLC VFAINNTKSF EDIHQYREQI
     EEEEEEE  S   TTS SSHH HHTT SEEEE EEETT HHHH HTHHHHHHHH

101  KRVKDSDDVP MVLVGNKCDL AARTVESRQA QDLARSYGIP YIETSAKTRQ
     HHHTTSS    EEEEEE TTS S  SS HHHH HHHHHHT   EEE  TTT T

151  GVEDAFYTLV REIRQHKLRK L
     THHHHHHHHH HHHHHHHHTT
```

Fig. 5.6 DSSP secondary structure alignments for **a** papain and **b** $p21^{ras}$. The key to the output is color-coded with the amino acids in that particular state. The meaning of the states listed in the key are discussed in the text. Compare the states with the types of helix listed in Table. 5.2 and notice the differences

Analysis of the assignments shows that they are very similar for papain, except for the beginnings or ends of the elements and the presence of a β-strand that is only found in the DSSP output. The same overall observation can be made from the comparison for $p21^{ras}$ (Table 5.4). For this protein, there is however one helix (A3) that appears to be interrupted in the DSSP analysis. We will have a look at this in the section on helices (▶ Sect. 5.7).

Table 5.3 Comparison of secondary structure assignments found in PDB, DSSP, and PyMOL outputs for papain. This analysis also shows that β-strands 1 and 3 of β-sheet 2, although consecutive in the primary sequence, are separated by β-strand 2

			PDB entry: 9pap	
		PyMOL	PDB	DSSP
α-helix				
L1		24–43	24–43	25–42
L2		50–57	50–57	50–56
L3		67–78	67–78	67–77
R1		117–128	117–128	117–127
R2		138–143	138–143	139–142
β-sheet 1				
Strand	1	5–7	5–7	5–6
	2	164–167	164–167	164–167
	3	170–175	170–175	170–174
	4	185–190	185–191	186–190
				148–149
β-sheet 2				
Strand	1	107–113	108–113	109–112
	2	206–210	206–210	207–210
	3	130–134	130–134	130–134
	4	158–163	158–163	158–163

Why do the assignments between the PDB and DSSP differ? I suspect that, as the structures of papain and $p21^{ras}$ were solved in 1986 and 1991, respectively, the authors themselves entered the secondary structure elements without using DSSP.

5.4.1 Secondary Structure Assignment by PyMOL

It is reassuring that the secondary structure assignments between PyMOL, the PDB, and DSSP are so similar because PyMOL employs a completely different method to the other two. PyMOL's method is based on the torsional angles of each C_α atom. We saw in ▶ Chap. 4 that, although the ϕ and ψ angles did repeat in secondary structure elements,

Table 5.4 Comparison of secondary structure assignments found in PDB, DSSP, and PyMOL outputs for p21^{ras}. This analysis also shows that β-strands 3 and 1, although consecutive in the primary sequence, are separated by β-strand 2

			PDB entry: 1q21	
		PyMOL	PDB	DSSP
α-helix				
A1		15–26	15–26	16–25
A2		69–75	69–75	69–72
A3		87–104	87–104	86–103 (res 92 turn)
A4		126–137	126–137	127–137
A5		151–171	151–171	152–168
β-sheet				
Strand	1	38–46	38–46	38–46
	2	49–57	49–57	48–57
	3	2–9	2–9	2–9
	4	77–84	77–84	77–83
	5	110–117	110–117	111–116
	6	140–144	140–144	141–143

the angles were not that regular. Thus, assigning secondary structure will not be so straightforward when it is based on the properties of the angles of the polypeptide backbone before and after the C_α atom. How does PyMOL use these angles for its assignments? It has a series of settings that start with "ss_helix" and "ss_strand" shown in Fig. 5.7. You can see these values for yourself in PyMOL by clicking on "setting" on the pull-down menu at the top and then on "edit all." Enter into the search box "ss_" and you will see the settings as in Fig. 5.7.

What do these settings mean? Let us take the settings for ss_helix_phi. PyMOL calculates the φ angle for a particular atom and compares it to the ss_helix_target_ value of −57.00°. Any ϕ angle within ±55.0° of −57.0° (ss_helix_phi_include) will be considered as a possible contender for an α-helix. Any ϕ values which are ±85.0° of −57.0° (ss_helix_phi_exclude) will be excluded. Depending on the adjacent residues in the polypeptide chain, values in between 55.0° and 85.0° can be considered as being in an α-helix. PyMOL assesses the ϕ and ψ angles in this way for the entire protein, includes or excludes them into helices and strands, and so assigns secondary structure. The deliberate "vagueness" of having the intermediate region between full inclusion and full exclusion reflects the difficulty of assigning secondary structure and allows the algorithm to be "flexible."

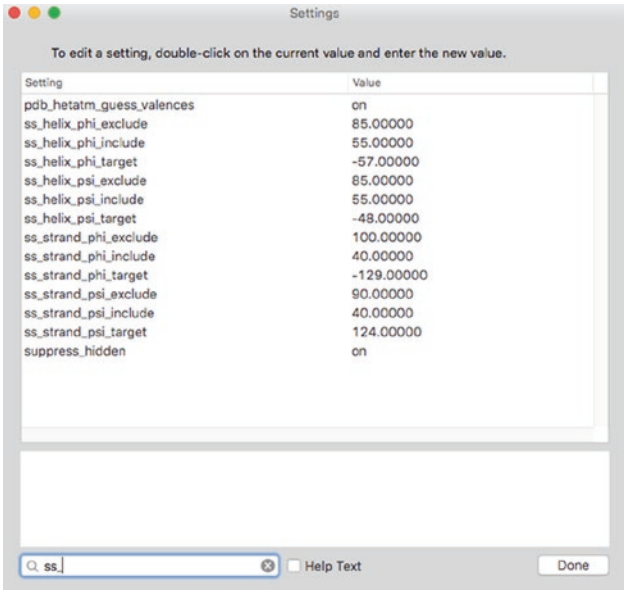


Fig. 5.7 Examining the settings for determining secondary structure assignments in PyMOL. These settings were brought up by entering "ss_" in the search box

It is important to always remember that algorithms are automation machines written by humans. Thus, they can be and are sometimes wrong, especially when a situation is an unusual one such as one that was not foreseen by the algorithm writer. If you wish, you can change these values yourself (e.g., "set ss_helix_phi_exclude, 70") and see what happens. Do not forget to reset the values (e.g., the command "reinitiate" or just "reini" for short) before starting something important.

Having considered how secondary structure is assigned, let us focus on these elements in more detail, starting with β-sheets. Subsequently, we will look at the turns between the β-strands in a β-sheet. As some of these turns have an arrangement that is almost helical, this will lead us into a discussion of α-helices.

5.5 Exploring β-Sheets

Let us start by examining the β-sheets in papain, shown as sticks in Fig. 5.8. The strand that is only assigned by DSSP (Table 5.3) has the carbon atoms in salmon; all the others are in magenta. ▶ Box 5.2 provides the instructions how to make the drawing.

Rotate your PyMOL drawing of Fig. 5.8 of the two β-sheets, viewing them both singly and together to explore their properties (e.g., orientations, directions, polar contacts, pleats, and twists). There are several points to note. First, notice that all the β-strands are anti-parallel, except for that assigned by DSSP only; β-sheet 1 is thus an anti-parallel β-sheet, whereas β-sheet 2 is a mixed one. Further, one β-strand has contiguous residues from both β-sheets (i.e., residues 158–163 and 164–167, see Table 5.3), thus conferring extra stability on the domain. Second, observe how the polar contacts between the amide hydrogens and carbonyl oxygen atoms of the peptide bonds occur at 90° to the direction

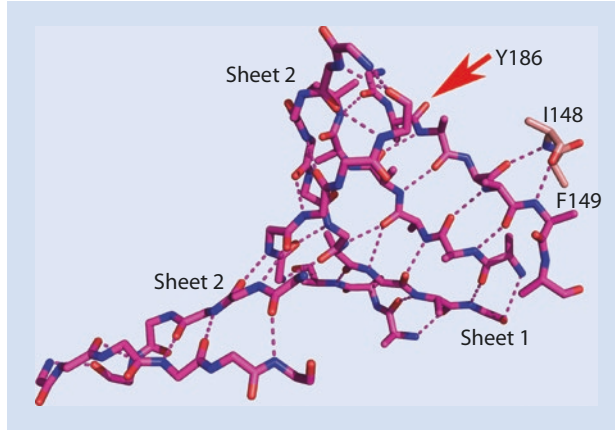


Fig. 5.8 The two β-sheets of papain. Stick representation of the main-chain atoms of papain in the two β-sheets. The atoms colored salmon are from residues 148 and 149 that were assigned as a β-strand by DSSP. Residue Tyr146 is referred to in the text. Note the twist of the sheets and their pleating

Box 5.2 Drawing and Altering the β-Sheets of Papain as in Fig. 5.8

```
#          Bring up the entry into PyMOL
Open the drawing papain_5_1a.pse
#          show the first β-sheet, including the extra strand assigned from DSSP
PyMOL>   create papain_sheet_1, 9pap and (i. 5-7 i. 164-167 i.
         170-175 i. 185-191 i. 148-149)
PyMOL>   hide everything, papain_sheet_1
PyMOL>   select papain_sheet_1
PyMOL>   show sticks, name n+ca+c+o and not hetatm and sele
GUI      sele, actions, find, polar contacts, just intra_main chain
GUI      sele_polar_conts, actions, rename, pol_pap_sheet_1
PyMOL>   util.cbam sele
PyMOL>   select papain_sheet_1 and name ca+cb
PyMOL>   show sticks, sele
#          color the atoms of the extra strand salmon
PyMOL>   select papain_sheet_1 and i. 148-149
PyMOL>   util.cbas sele
#          show the second β-sheet
PyMOL>   create papain_sheet_2, 9pap and (i. 108-113 i. 206-210 i.
         130-134 i. 158-163)
PyMOL>   hide everything, papain_sheet_2
PyMOL>   select papain_sheet_2
PyMOL>   show sticks, name n+ca+c+o and not hetatm and sele
PyMOL>   util.cbam sele
GUI      sele, actions, find, polar contacts, just intra_main chain
GUI      sele_polar_conts, actions, rename, pol_pap_sheet_2
PyMOL>   select papain_sheet_2 and name ca+cb
PyMOL>   show sticks, sele
PyMOL>   set dash_color, magenta
#          add to the cartoon the extra β-strand assigned by DSSP
PyMOL>   alter 9pap and i. 148-149, ss='S'
PyMOL>   rebuild
PyMOL>   select 9pap and i. 148-149
PyMOL>   color salmon, sele
#          change the background color
PyMOL>   bg white
#          to make the drawing in Fig. 5.8, use the mouse to hide the ribbon by clicking on
         9pap on the GUI on the right
#          Save the session
> File, Save Session as "papain_5_8", Quit
```

of the polypeptide chain. Are all of the polar contacts indicated by PyMOL really hydrogen bonds? For β-sheet 1, this is the case. For β-sheet 2, there is one polar contact that cannot be a hydrogen atom. Can you find it? (hint: think about the number of hydrogen bonds that can be made by the carbonyl oxygen and amide nitrogen atoms)

A third important aspect to be gleaned from the examination of the polypeptide backbones of the β-strands is that they are not completely extended. This can be seen by measuring the ϕ and ψ angles for residues in the strands. Choose some of the residues and measure their torsion angles. You will observe that most of the values for the torsion angles are closer to −135° for ϕ and 135° for ψ than to values of −180° and 180° that are required for a fully extended chain. Leu134 is one of the residues in papain in a β-strand that has ϕ and ψ values closest to −180° and 180°. Not surprisingly, the observed values of the torsion angles fit well with the target values shown above in the PyMOL algorithm for defining sheets given in ◘ Fig. 5.7. If you use the VADAR web page mentioned in ▶ Chap. 4 for 9pap, you can see all of the ϕ and ψ angles for the protein. Look at those for the β-strands (marked B in the fourth column from the left) to confirm that the values are mostly around −135° for ϕ and 135° for ψ and that these values do indeed repeat. Look also for residues with ϕ and ψ angles that are close to −180° and 180°. What do you notice? There are several that are not in β-strands. Which type of residue are they? Look using PyMOL at their location in papain to see their effect on the structure.

The absence of the complete extension of the polypeptide chain also means that the distance between the C_α atoms from residue i and residue $i + 2$ (e.g., from the C_α atoms of residues 132 and 134) is less than the 7.6 Å theoretically predicted from the bond lengths. Thus, even the distances between the C_α atoms of residues 132 and 134 and between 133 and 135 are only 7.0 Å and 7.2 Å, respectively, even though Leu134 is well extended.

Measuring the distances between residues i and $i + 2$ in strands also reveals that the strand can be considered a repeating unit of two amino acids. In other words, if the chain is moving to the right on an imaginary x axis, the C_α atoms of residues i and $i + 2$ are at the same position on an imaginary y axis. Furthermore, the C_β atoms on these C_α atoms point in the same direction. If these C_β atoms are pointing upward, the bonds from their C_α atom to the carbonyl carbon atom must point down due to the tetrahedral geometry of the carbon atom. Consequently, there is an up and down arrangement of the polypeptide chain that gives each strand a "pleated" appearance. The term "pleat" originates from the clothing industry. If you enter "kilt pleat" into a search engine of your choice, you will find plenty of images.

Why are values of −135° for ϕ and 135° for ψ preferred in β-sheets? Why is the polypeptide chain not fully extended? When the backbone takes these preferred values, the side-chains lie almost at right angles to the chain, so that there is little steric clash of the C_β atom of the side-chain with the atoms of the backbone (e.g., with the carbonyl oxygen atom of the same residue). In other words, the ϕ and ψ values of −135° and 135° are energetically more favorable than those found for the fully extended polypeptide chain. Furthermore, the ϕ and ψ angles found in β-sheets are found in an area of the Ramachandran plot with the lowest steric clashes for the C_α atoms. Thus, the torsion angles of the β-strands can utilize the considerable space in this area of the plot, giving the β-strands a variability that is illustrated in ▶ Chap. 8.

One aspect of β-structures that we have not examined is the β-bridge that was assigned by the DSSP algorithm (◘ Fig. 5.6). This is defined by Kabsch and Sander (1983a) as one residue with ϕ and ψ angles in the β-sheet area of the Ramachandran plot that form just two hydrogen bonds with another residue. An example is residue Glu183 in papain which forms a β-bridge with Tyr186; Tyr186 is itself at the edge of β-sheet S1 in papain (◘ Fig. 5.8). If you look at the drawing you made of papain corresponding to ◘ Fig. 5.8, you can create a new object for Glu183 and Tyr186 (e.g., create pap_183_186, 9pap and i. 183–186) and ask PyMOL to show you the hydrogen bonds by finding polar contacts.

In the terminology of Kabsch and Sander, the β-bridge is considered to be the basic building block of the β-sheet. When two β-bridges are consecutive, this is termed a ladder of two anti-parallel strands; when two ladders are consecutive, a β-sheet is formed.

As the region of papain from residues Gln178 to Tyr186 is an illuminating, if complex, one, we will return to these residues in the section on β-turns.

5.5.1 Looking into the β-Strand Assigned Only by DSSP

We have investigated the anti-parallel β-sheets of papain. As an introduction to the parallel β-sheets of $p21^{ras}$, we can look to see why PyMOL does not find the short parallel β-strand that was assigned by DSSP (Table 5.3). To investigate this question, zoom in on residues 148–149 in the papain drawing in Fig. 5.8 made from ▶ Box 5.2 and measure the distances between the following pairs of atoms: carbonyl oxygen of Gly147 and amide nitrogen of Arg188, amide nitrogen of Phe149 and carbonyl oxygen of Arg188, and carbonyl oxygen of Phe149 and amide nitrogen of Lys190. The first two atoms are too distant for a hydrogen bond whereas the latter two are close enough to form a hydrogen bond. Thus, only two hydrogen bonds can be formed. Center the drawing on residue 188 (center i. 188) to compare the arrangement of the backbones. Note the difference in the arrangements of the parallel and anti-parallel strands in terms of the amide nitrogen atoms and carbonyl oxygen atoms. The carbonyl oxygens in the anti-parallel strands point directly at the amide nitrogen; in the parallel strand, they are at an angle.

Returning to the assignment of the β-strand, measure the ϕ and ψ angles for residues 147, 148, and 149. The ϕ angles for 148 and 149 lie outside the "include" limits for PyMOL for immediately assigning a β-strand but inside the "fuzzy area" that depends on the local environment (Fig. 5.7). Presumably, these measurements, together with the presence of just two hydrogen bonds in a parallel sheet, preclude PyMOL from assigning a β-strand at this position. DSSP assigns a β-strand here because Phe149 hydrogen bonds to residues 188 and 189, that is, two consecutive β-bridges are present.

5.5.2 Investigating Parallel β-Strands

The secondary structure of $p21^{ras}$ is shown in Fig. 5.9a. We saw earlier that it contains one mixed β-sheet that comprises one anti-parallel strand and five parallel ones (Fig. 5.9b). Instructions for preparing Fig. 5.9 are in ▶ Box 5.3.

The mixed nature of the β-sheet permits an exact comparison of the two anti-parallel strands 1 and 2 with the parallel ones. Note that β-strand 2 is anti-parallel to strand 1 but parallel to strand 3. Compare the orientations of the polar contacts between strands 1 and 2 and between 2 and 3. In the anti-parallel strands, there is always just one C_α between the pairs of polar contacts. In contrast, in the parallel strands (2 and 3), there are always three atoms (i.e., one entire residue) between the hydrogen bonds on one of the strands. For this reason, the polar contacts are not at 90° to the polypeptide chain and the parallel β-sheet is weaker than the anti-parallel one.

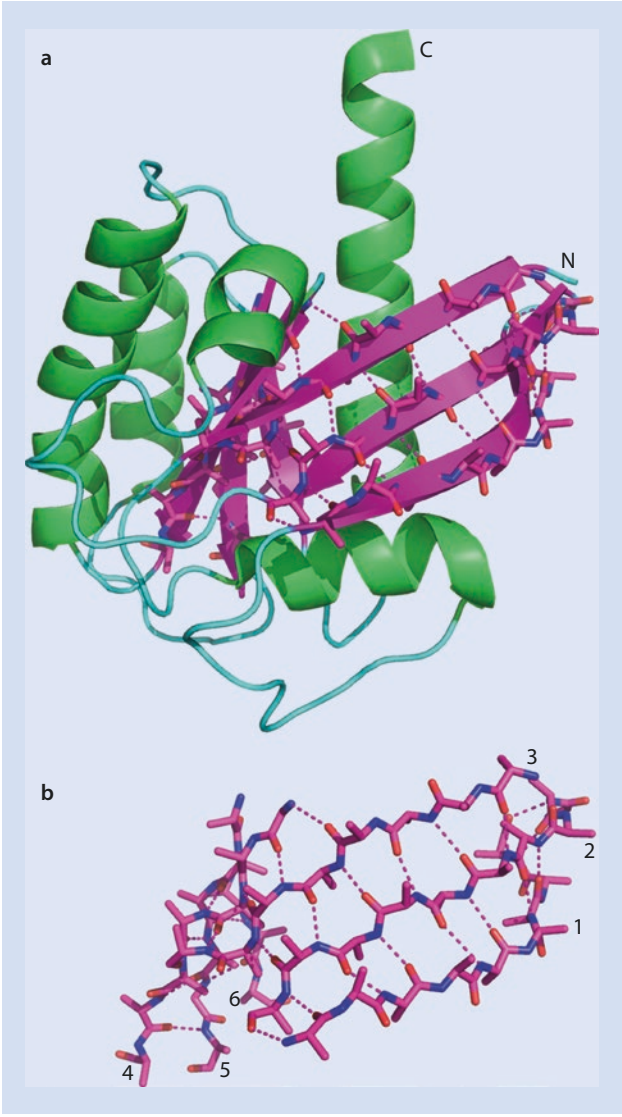


Fig. 5.9 The secondary structure of p21^{ras}. **a** Cartoon representation with the β-sheet backbone and the C_β atoms also shown as sticks. **b** View of just the backbone and C_β atoms showing the mixed β-sheet. The bottom two β-strands are anti-parallel, the remainder are parallel. Note the differences in the direction of the hydrogen bonds and the twist of the β-strands as shown by the positions of the C_β atoms

You may have noticed that the β-sheets in papain and p21^{ras} are all twisted to the right. The exact reason for this twisting is still the subject of debate (► http://kinemage.biochem.duke.edu/teaching/anatax/html/anatax.2b.html); the ϕ and ψ torsion angles characteristic of the β-sheet together with a distortion of the geometry at the amide nitrogen atoms may be a contributing factor (Weatherford and Salemme 1979). The right-hand twist is easily visualized by the cartoon in Fig. 5.9a and by the positions of the C_β atoms in the drawings with sticks (Fig. 5.9b).

We will look further at the topology of p21^{ras} in ► Chap. 6 (► Sect. 6.3). We now turn to the structure of the turns that cause the chain to move in a different direction.

Box 5.3 Drawing the β-Sheets' Main-Chain Polar Contacts of p21^{ras} as in Fig. 5.9

```
#           bring up the entry into PyMOL
PyMOL>      fetch 1q21
asPyMOL>    as cartoon
#           Color according to secondary structure (ss)
PyMOL>      color green, ss h
PyMOL>      color magenta, ss s
PyMOL>      color cyan, ss l+
Mouse>      rotate the molecule so that the β-sheet domain is at the
            front and the long C-terminal α-helix is at the back on
            the right
#           show the sequence of the entry colored according to secondary structure
PyMOL>      set seq_view, on
#           show the β-sheet using the DSSP assignments
PyMOL>      create ras_sheet, 1q21 and (i. 38-46 i. 48-57 i. 2-9 i.
            77-83 i. 111-116 i. 141-143)
PyMOL>      select ras_sheet
PyMOL>      show sticks, name n+ca+c+o and sele
PyMOL>      hide cartoon, ras_sheet
PyMOL>      util.cbam sele
#           add the Cβ-atoms to see the side-chain positions
PyMOL>      select ras_sheet and name ca+cb
PyMOL>      show sticks, sele
GUI         ras_sheet, actions, find, polar contacts, just intra_main
            chain
GUI         ras_sheet_polar_conts, actions, rename, pol_ras_sheet
PyMOL>      set dash_color, magenta
#           change the background color
PyMOL>      bg white
#           to view just one of the objects, click on the object name in the GUI at the top
            right
#           save the session
> File, Save Session as "ras_sse_5_9", Quit
```

5.6 Turns

5.6.1 β-Turns

The examination of the β-strands in this chapter illustrated that the polypeptide chain must change its direction to allow consecutive β-strands to be able to hydrogen bond to each other. Often, these turns of direction can be quite sharp, as can be seen between residues Gly167 and Tyr170 in papain and Ile46 and Glu49 in p21^{ras}. These residues are shown in Fig. 5.10, the drawing instructions are in ▶ Box 5.4, and the φ and ψ angles measured are shown in Table 5.5. You may find that the turn formed by residues 167–170 is surprising. Why do I suggest this? First, Gly167 is not part of the turn but is still part of the β-strand. It is indeed, like Tyr170, quite extended; can you see this extended nature without determining the torsional angles? (Hint: look at the positions of the amide nitrogens and carbonyl oxygens.) Second, it is Pro168 and Asn169 that turn the polypeptide backbone; for each of the residues, the chain turns 90°.

There are several types of turns found in proteins (Hutchinson and Thornton 1994). They are characterized by their hydrogen-bonding arrangements and their φ and ψ angles.

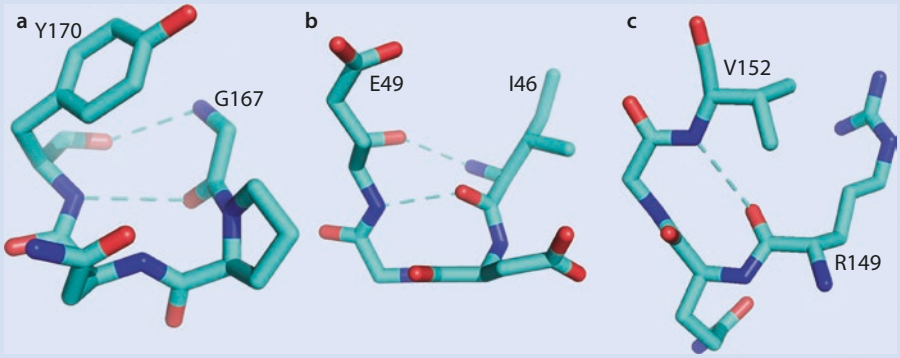


Fig. 5.10 Three types of turn. **a** Type I β-turn formed by residues Gly167–Tyr170 of papain. **b** Type I′ β-turn formed by residues Ile46–Glu49 of $p21^{ras}$ **c**. The type II β-turn formed by residues Arg149–Val152 of $p21^{ras}$

Box 5.4 Drawing the Type I β-Turns in Fig. 5.10

```
#          A, Gly167 and Tyr170 of papain
#          bring up the entry into PyMOL
PyMOL>     fetch 9pap
PyMOL>     hide everything
PyMOL>     select i. 167-170
PyMOL>     show sticks, sele
PyMOL>     zoom sele
Mouse>     rotate the molecule to have the orientation in Fig. 5.10a
PyMOL>     util.cbac sele
GUI        sele, actions, find, polar contacts, just intra_main chain
GUI        sele_polar_conts, actions, rename, pol_pap_167-170
PyMOL>     set dash_color, cyan
PyMOL>     set dash_length, 0.4
#          change the background color
PyMOL>     bg white
#          save the session
>File, Save as "turn_167_170_5_10a", Quit
#          B, Ile46-Glu49 of p21ras
#          bring up the entry into PyMOL
PyMOL>     fetch 1q21
PyMOL>     hide everything
PyMOL>     select i. 46-49
PyMOL>     show sticks, sele
PyMOL>     zoom sele
Mouse>     rotate the molecule to have the orientation in Fig. 5.10b
PyMOL>     util.cbac sele
GUI        sele, actions, find polar contacts, just intra_main chain
GUI        sele_polar_conts, actions, rename, pol_p21_46-49
PyMOL>     set dash_color, cyan
PyMOL>     set dash_length, 0.4
#          save the session
>File, Save as "turn_46-49_5_10b"
```

```
#          C, Arg149-Val152 of p21ras
PyMOL>   hide everything
PyMOL>   select i. 149-152
PyMOL>   show sticks, sele
PyMOL>   zoom sele
Mouse>   rotate the molecule to have the orientation in Fig. 5.10c
PyMOL>   util.cbac sele
GUI      sele, actions, find, polar contacts, just intra_main chain
GUI      sele_polar_conts, actions, rename, pol_p21_149-152
#          save the session
> File, Save Session as "turn_149-152_5_10c", Quit
```

5

Table 5.5 φ and ψ angles measured in the indicated types of turns

Drawing	Protein	Turn/ type	Torsion angle (°)							
5.10			Position							
			i		*i* + 1		*i* + 2		*i* + 3	
			ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ
	Papain 167–170	βI	−149.5	−172.3	−54.1	40.2	−124.7	11.5	148.0	165.2
	p21ras 46–49	βI′	−122.3	115.9	54.3	47.6	77.1	−19.3	142.6	132.9
	p21ras 149–152	βII	79.5	6.6	−62.0	127.8	80.1	−13.2	−60.2	-41.3
5.11	Papain 22–24	γ′	−140.6	132.5	−91.4	69.1	102.7	−12.9	n.a	
	Papain 199–202	3_{10}	−66.0	130.7	56.0	34.5	60.5	34.1	−60.9	−20.9

n.a. not applicable

The three turns illustrated in Fig. 5.10 have a hydrogen-bonding arrangement termed *i* to *i* + 3, indicated by the polar contacts in cyan in the figure. In other words, the residue (termed *i*) accepting the hydrogen is residue *i*; the residue donating the hydrogen atom is three residues distant and thus termed *i* + 3. This bonding pattern, together with their φ and ψ angles, classifies this turn of papain as a type I β-turn and that of p21ras as a type I′ β-turn (often termed an inverse β-turn). The relationship between the two is that the φ and ψ angles of the inverse turn have the opposite signs to those in the standard turn (Table 5.5); consequently, the main-chain atoms have the opposite chirality. Type I and I′ β-turns are very common in proteins, as are type II and II' β-turns. p21ras has an example

of a type II turn from residues Arg149–Val152, as shown in Fig. 5.10c, lying between an α-helix and a β-strand. It can be drawn in the same way as in ▶ Box 5.4 by replacing the indicated amino acids.

Compare the arrangement of the three turns illustrated here in Fig. 5.10. In two out of three cases, the amino acid at $i + 2$ is glycine. This is not a statistical blip. Glycine is very often found at this position in type I′ as well as type II and II' β-turns. Why should this be? Imagine another amino acid at this position instead of glycine. In the type II β-turn, the C_β of the side-chain would clash with the carbonyl oxygen. In the type I′ β-turn, the clash would be with the C_β and the amide hydrogen.

Where can I find out about these classifications and the designations of turns in general? Where are these classifications for a particular protein documented? Information on the turns themselves are of course cited in research publications such as classic ones by Janet Thornton and her colleagues (Wilmot and Thornton 1988, 1990; Hutchinson and Thornton 1994). Indeed, these papers provided many of the bases for these classifications. However, an excellent source of much detailed information is available in PDBsum, the website (▶ www.ebi.ac.uk/pdbsum) that we used to look at Ramachandran angles in ▶ Chap. 4. Let us look at the turns in papain and $p21^{ras}$ using this database; the assignments in PDBsum are made by an algorithm termed PROMOTIF (Hutchinson and Thornton 1996). Unfortunately, this program is not available online.

To see the information in PDBsum on turns, enter the PDB identifier into the search window as in ▶ Chap. 4 and then select "protein" from the tab list above the entry. Here, we see a wiring diagram similar to that in the PDB as well as a red topology diagram that we will look at in ▶ Chap. 6. Now, we need to look to the bottom left for the list of motifs found in the protein. For $p21^{ras}$, we note that there are 14 β-turns. Click on this item to find their location. On the subsequent page, all of the β-turns are fully listed with their types and their ϕ and ψ angles (the χ_1 angle of the C_α atom to the C_β atom is also listed). At the top, the turns are shown graphically. Information on the plots, the data in the table, and the nine classes of turn can be found by clicking the link to motif description. The type I and type II turns are the most common β-turns; as they reverse the direction of the chain, they are also termed reverse turns. For this and for reasons of space, we have only considered these types of β-turns in this book.

5.6.2 γ-Turns

A γ-turn is a tighter one than the β-turn as it results from a hydrogen-bonding pattern of residue i to $i + 2$ and has ϕ and ψ angles within 40° of 75.0° and −64.0°, respectively, for a classic γ-turn (Rose et al. 1985, Milner-White et al. 1988). An inverse γ-turn has again the opposite signs on the torsion angles. There are none in the $p21^{ras}$ structure; however, there are four inverse γ-turns in the papain structure. The γ-turn formed by residues Cys22 to Ser24 in papain is shown in Fig. 5.11; it can be drawn using the PyMOL commands that you have learnt so far. The positions of the neighboring residues in the drawing illustrate that the γ-turn produces a kink in the polypeptide chain rather than an actual turn. As with the β-turn, PDBsum provides a wealth of information on these turns in general and on those in papain itself. Have a look at the other three and their positions in the papain structure.

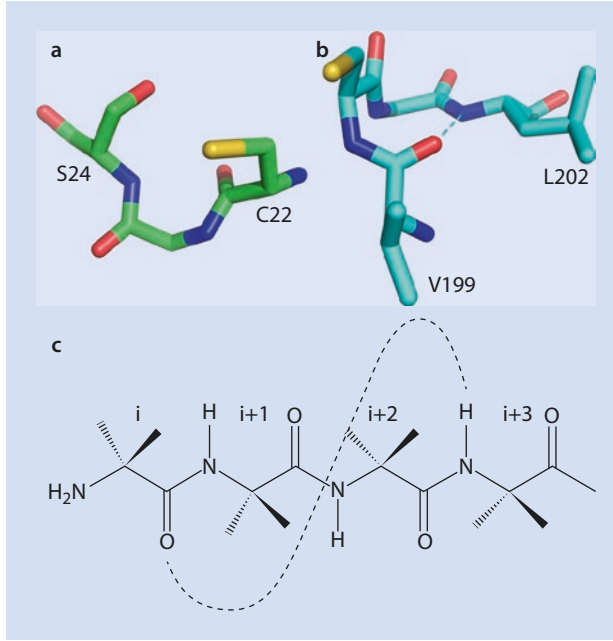


Fig. 5.11 A γ-turn and a helical 3_{10} turn. **a** γ-turn formed by residues Cys22-Ser24 in papain. Note that the change in the direction of the backbone is more like a kink than a turn. Why does PyMOL not show a polar interaction between the carbonyl oxygen of Cys22and the amide nitrogen of Ser24? **b** The 3_{10} turn formed by residues Val199 to Leu202 in papain. **c** The 10-atom hydrogen-bonded ring in the 3_{10} turn

5

5.6.3 The Helical 3_{10} Turn (or a Type III β-Turn)

In the list of turns given by PDBsum, you may have noticed that there is no definition for a type III β-turn. The reason is that β-turns previously classified as type III have the same torsional angles ($\phi = -60°$, $\psi = -30°$) as a type of helix termed a 3_{10} helix. The parameter 3 means that there three residues per turn; the parameter 10 indicates that there are ten atoms in the bonded ring (Fig. 5.11c). As an example, we can look at residues Val199 to Leu201 in papain. This region is assigned as a 3_{10} helix both by DSSP (Fig. 5.6, marked with G) and by PDBsum. To find the turn in PDBsum, we need to click on helices instead of turns. Again, the 3_{10} helix is marked with a G. Figure 5.11 (again, you should be able to draw this yourself) shows this 3_{10} helix. If you count the atoms in the hydrogen-bonding ring and only find nine, you have probably forgotten the hydrogen. If we look at the torsion angles for Cys200 and Gly201, we find values of 56.0° and 34.5° for Cys200 and 60.5° and 34.1° for Gly201. Normally, the values for a 3_{10} helix are around −60° and −30°. The example here is thus one in which the inverse angles are present.

We will see also later in the chapter that 3_{10} helices can be found at the end of α-helices.

5.6.4 The α-Turn with Bonding from Residue *i* to *i* + 4

For the sake of completion, the α-turn with bonding from residue *i* to *i* + 4 should also be mentioned. They are not so common as the other types of turn, but they are possible nevertheless. The bonding, as we will see below, is the same as in α-helices; however, the ϕ and ψ angles are different. The turns have been described and collated in detail by Pavone et al. (1996); indeed, they suggest in this paper that residues Asp6 to Lys10 of papain form such an α-turn. DSSP assigns this region to be a hydrogen-bonded turn. If you examine

this region, you can indeed find the hydrogen-bonding pattern from the carbonyl oxygen of residue *i* to the amide nitrogen of residue *i* + 4; it is however obviously not an α-helix.

5.6.5 The Complexity of Turns in Proteins

We have seen that turns are classified by their ϕ and ψ angles and hydrogen-bonding patterns. You should have noticed that, in addition to glycine and proline, small amino acids such as asparagine, aspartic acid, and serine are common in turns. These residues can take on unusual torsional angles and, in the case of asparagine and aspartic acid, even hydrogen bond back to the polypeptide chain.

As you have perhaps realized, we have only scratched the surface of the subject of turns in proteins. Given that turns comprise between 25% and 45% of protein structures (Chou and Fasman 1977), this is perhaps not surprising. To illustrate how complex turns can be, look at residues Gly178 to Tyr 186 in papain in PyMOL, in DSSP (◘ Fig. 5.6), and in PDBsum. We looked at Glu183 because it is a β-bridge. However, as you can see from PDBsum or DSSP, several of the other residues are involved in β-turns or in bends. Such a region is not surprisingly termed a multiple turn (Hutchinson and Thornton 1994). They are found quite frequently in proteins.

If you are interested in finding out more about turns, you will find the following references useful: Chou and Fasman (1977), Rose et al. (1985), and Panasik et al. (2005).

5.7 Exploring Helical Structures of Proteins

5.7.1 α-Helices

We saw in the previous section the α-turn with the hydrogen bond pattern from residue *i* to *i* + 4 but without the characteristic ϕ and ψ angles of the α-helix. Let us look now at this bonding pattern in the α-helix with its characteristic ϕ and ψ angles in the region of −60° and −40°. As examples of α-helices, we will examine residues Thr87 to Val103 from 1q21. In ◘ Fig. 5.4, these residues are assigned in the PDB as one α-helix whereas DSSP assigns all residues as an α-helix except for Asp92 which it designates as a turn (◘ Fig. 5.6). The residues are shown in ◘ Fig. 5.12a as a cartoon and backbone. In panel B, the residues are just shown as backbone; in panel C, the residues around Asp92 are viewed in close-up. The C_β atoms are also visible in each of the drawings. ▶ Box 5.5 explains how to draw it.

Examination of ◘ Fig. 5.12 shows indeed that residues 87–103 actually build two helices as depicted by DSSP and PDBsum. We will also see below that the helical parameters of the two helices vary. First, though, let us observe the fundamental properties that are common to both. First, note the orientation of the amide nitrogens and the carbonyl oxygens. If the polypeptide chain runs from the top (as in ◘ Fig. 5.12), the amide nitrogens (and thus the amide hydrogens) will point up and the carbonyl oxygens will point down. The C_β atoms of the side-chains, lying on the outside of the helix, will point upward at an angle of about 45° from the horizontal. The presence of the C_β atoms aids you in noticing that the α-helix is right-handed. To see this best, look up the helix from the bottom in your PyMOL drawing from ◘ Fig. 5.12 and note the stagger of the C_β atoms. To observe the right-handedness, hold your right thumb up and rotate it in the same direction as the C_β stagger. Your thumb should turn to the right.

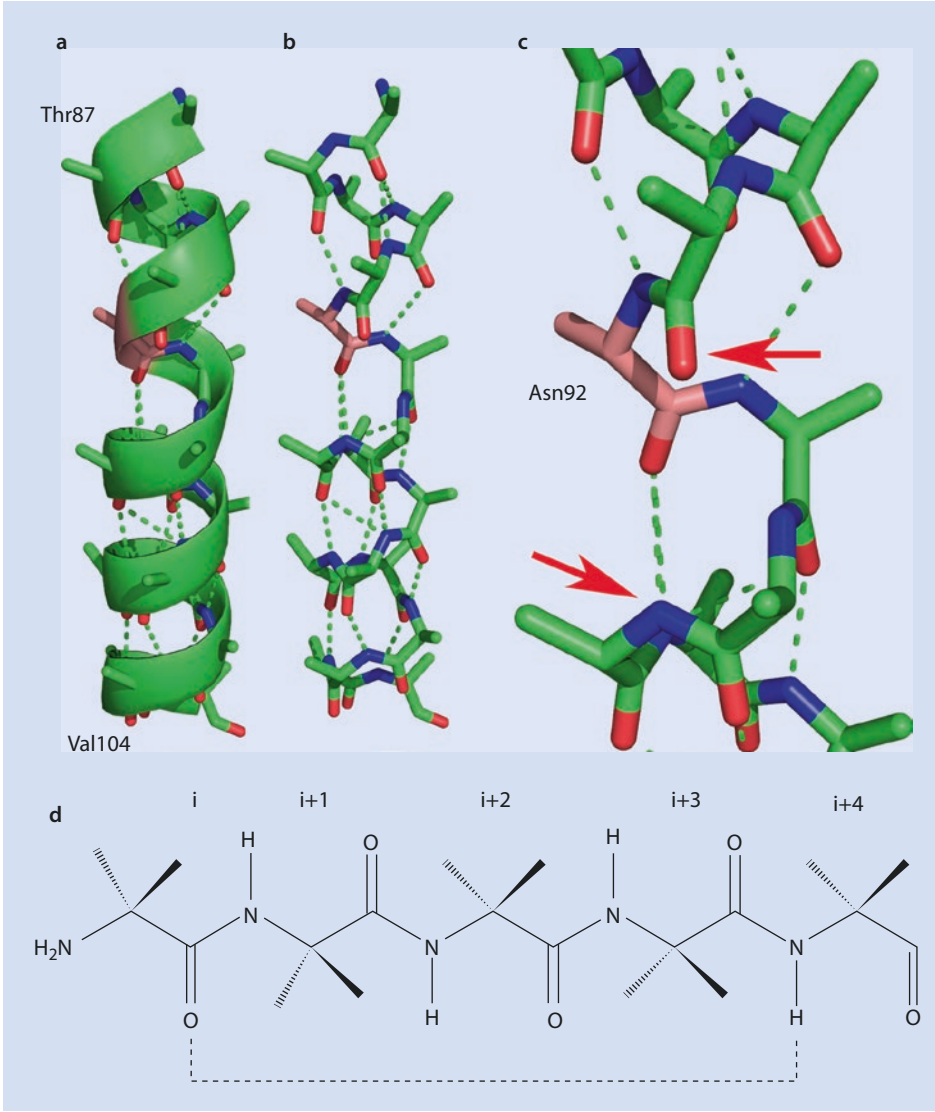


Fig. 5.12 PyMOL images of the α-helix built by residues Thr87 to Val104 in $p21^{ras}$. **a** Cartoon and stick representation. **b** Stick representation of backbone and C_β atoms. **c** Enlargement of residues Ala90 to His94. Carbon atoms are green, except for those of Asp92 which are salmon. Oxygen atoms are red, nitrogen blue. The arrows indicate that the carbonyl oxygen and amide hydrogen are not close enough for a polar contact, leading to an interruption of the α-helix. **d**. The 13-atom hydrogen-bonded ring in the α-helix

If you count the atoms in your PyMOL drawing (Fig. 5.12), you can confirm the hydrogen-bonding pattern to be from the carbonyl oxygen of residue *i* to the amide nitrogen of residue *i* + 4 with 13 atoms in the hydrogen-bonded ring. Once again, there are more polar contacts shown by PyMOL than hydrogen bonds, as the amide hydrogens can only undergo one hydrogen bond. Remember that the hydrogen bonds that are straight are the strongest. For these to be straight between the carbonyl oxygens and the amide nitrogens, they must be aligned above each other. Thus, the helix must repeat every 3.6

Box 5.5 Drawing Residues Thr87 to Val104 of p21^{ras} as in ◘ Fig. 5.12

```
#          Bring up the entry into PyMOL
PyMOL>     fetch 1q21
PyMOL>     as cartoon
Mouse>     Rotate the molecule so that the β-sheet domain is at the
           front and the long C-terminal α-helix is at the back on
           the right
#          show the sequence of the entry
PyMOL>     set seq_view, on
#          make separate object for the helix
PyMOL>     create helix, i. 87-104
PyMOL>     hide everything, 1q21
PyMOL>     select helix
PyMOL>     show sticks, helix and name n+ca+c+o
PyMOL>     hide cartoon
PyMOL>     util.cbag sele
#          add the Cβ-atoms to see the side-chain positions
PyMOL>     select helix and name ca+cb
#          A PyMOL short-cut. The program just needs just an unambiguous shortened
           command to perform an operation
PyMOL>     sh sti, sele
#          color residue Asp92 salmon
PyMOL>     select i. 92
PyMOL>     util.cbas sele
#          show polar contacts in the α-helix
PyMOL>     select helix and name n+ca+c+o
GUI        sele, actions, find, polar contacts, just intra_main
           chain
GUI        sele_polar_conts, actions, rename, pol_ras_ helix
PyMOL>     set dash_color, green
Mouse>     rotate the molecule 180° so that the polypeptide chain
           of the α-helix runs from top to bottom
#          change the background color
PyMOL>     bg white
#          Save the session
> File, Save Session as "ras_helix_5_12", Quit
```

residues (carbonyl carbon atom from residue *i*; all three backbone atoms from the three residues $i + 1$, $i + 2$, and $i + 3$; and the amide nitrogen atom from $i + 4$). At the N-terminus of the helix, the amide hydrogens of the first three to four residues are not hydrogen bonded in the *i* to $i + 4$ pattern because there are no carbonyl oxygens directly above them; at the C-terminus, there are three to four carbonyl oxygens that do not have binding partners in the *i* to $i + 4$ pattern because there are no amide hydrogens directly below them.

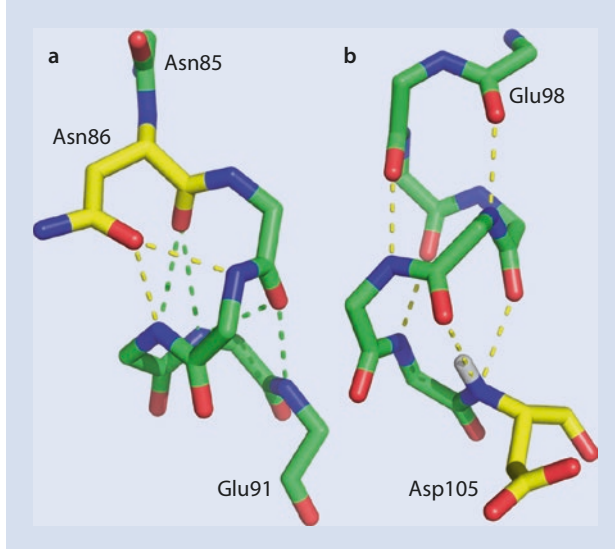
The α-helix is often referred to as the 3.6_{13} helix to indicate the number of atoms in the bonded ring and the number of residues per turn (◘ Fig. 5.12d). A further characteristic of the α-helix is, of course, the repeating ϕ and ψ angles. Measure the ϕ and ψ angles for residues His94–Glu98. You should find values of around −60° for ϕ and −40° for ψ. They are much less variable than those found in the β-strands because of the greater potential for stereochemical clashes in an α-helical structure. Nevertheless, the α-helices in a protein do deviate from the theoretically predicted optimal parameters. Before we investigate the deviations in more detail, let us examine the torsional angles of Asp92, the salmon-colored residue in ◘ Fig. 5.12. When you measure them, you should find values of −67.8° and −9.9°

for ϕ and ψ, respectively. As a consequence, the C_β atom of the side-chain and the carbonyl oxygen are almost in *trans*, an unusual state of affairs, and the α-helical structure is broken by this residue. This can also be seen when examining the hydrogen bonding from the drawing in ◘ Fig. 5.12c; because of the non-α-helical torsional angles of Asp92, there is no hydrogen bond between Glu91(i) and Gln95 (i + 4), as indicated by the red arrows. Asp92 is, however, in a position to correctly hydrogen bond to Lys88 and Tyr96.

Returning to the variations in helices, we can once again employ PDBsum to examine this point. Go to the "protein" page of the PDBsum entry for 1q21 and click on "helices." The first thing we note is the presence of six helices as suggested by DSSP and not five as suggested by PyMOL and PDB (◘ Table 5.4). Like DSSP, PDBsum has Thr87–Glu91 as one helix and Ile93–Val103 as another. The variation in the α-helices is given the column entitled "deviation from ideal" second from the right in table of properties. All of the α-helices in $p21^{ras}$ deviate from an ideal helix to a certain extent with the α-helix from Thr87 to Glu91 being one of the most deviant and that from Ile93 to Val103 one of the least. The deviations can be seen by looking at the pitch (i.e., the rise of the α-helix in Å for one turn) and the number of residues for one turn. For an idealized α-helix, these values are 5.4 Å and 3.6 residues. For the helix Thr87–Glu91, the corresponding values are 5.16 Å and 3.56; in contrast, the values for Ile93–Val103 are 5.70 Å and 3.66.

5.7.2 Variations Found at the N- and C-Termini of α-Helices

I mentioned earlier that the amide hydrogens at the N-terminus of an α-helix and the carbonyl oxygens at the C-terminus are not hydrogen bonded in the i to i + 4 pattern. Often, however, these atoms do form hydrogen bonds with atoms from other residues to form structures that are found in many α-helices. These structures are termed N- and C-terminal caps (Richardson and Richardson 1988). N-terminal caps very often involve a small residue that can hydrogen bond with an amide hydrogen at the top of the helix. Asn86 plays such a role in $p21^{ras}$ as shown in ◘ Fig. 5.13a. Commands are in ▶ Box 5.6.



◘ Fig. 5.13 N- and C-terminal caps of α-helices. **a** N-terminal cap in $p21^{ras}$ formed by residues Asn85–Glu91. **b** C-terminal cap formed by residues Glu98–Asp105. The carbon atoms of residues Asn86 and Asp105 are in yellow; all other carbon atoms are in green. Oxygen atoms are in red, nitrogen in blue, and the amide hydrogen of Asp105 in gray. The interactions between the side-chains of Asn86, Asp105, and the α-helix are shown in yellow. All other polar contacts are in green

Box 5.6 Drawing the N- and C-Terminal Caps of an α-Helix as in Fig. 5.13

```
#          Bring up the entry into PyMOL
PyMOL>     fetch 1q21
PyMOL>     hide everything
#          show the sequence of the entry
PyMOL>     set seq_view, on
#          N-terminal cap
PyMOL>     create n_cap, i. 85-91
PyMOL>     show sticks, n_cap and name n+ca+c+o
PyMOL>     util.cbag n_cap and name n+ca+c+o
PyMOL>     show sticks, n_cap and i. 86
PyMOL>     util.cbay n_cap and i. 86
GUI        n_cap, actions, find polar contacts, just intra_main
           chain
GUI        ncap_polar_conts, actions, rename, pol_n_cap
PyMOL>     select n_cap and i. 86 and name od1
GUI        sele, actions, find polar contacts, to any atoms in
           object
GUI        sele_polar_conts, actions, rename, pol_i_86
PyMOL>     color green, pol_n_cap
Mouse>     Rotate the drawing to the view in Fig. 5.13a
#          C-terminal cap
PyMOL>     create c_cap, i. 98-105
PyMOL>     show sticks, c_cap and name n+ca+c+o
PyMOL>     util.cbag c_cap and name n+ca+c+o
PyMOL>     sho sti, c_cap and i. 105
PyMOL>     util.cbay c_cap and i. 105
GUI        c_cap, actions, find polar contacts, just intra_main
           chain
GUI        c_cap_polar_conts, actions, rename, pol_c_cap
#          show the hydrogen atom to see the i + 3 bonding pattern
PyMOL>     h_add c_cap and i. 105 and name n
Mouse>     Rotate the drawing to the view in Fig. 5.13b
PyMOL>     bg white
#          Use "zoom n_cap" or "zoom c_cap" to jump between the drawings
#          Save the session
> File, Save Session as "n_c_cap _5_13", Quit
```

Indeed, if you count the atoms between the amide oxygen of Asn86 and the amide nitrogen of the Ser89, you will find 13 residues in the hydrogen-bonded ring. Measure the distance between the oxygen and nitrogen atoms; it is only 2.7 Å. It is as if the side-chain of Asn86 is the initiating point for the α-helix.

The situation at the C-terminus is different. If you look at the C-terminus of the α-helix in Fig. 5.12b, you may observe that the final turn does not move downward as much as the previous turns (in other words, the pitch is shorter). In Fig. 5.13b, this region is enlarged so that this effect is more evident. It can also be seen that the hydrogen-bonding pattern has changed so that the amide hydrogen (shown exceptionally in Fig. 5.13b) of Asp105 is now hydrogen bonding to the carbonyl oxygen of Arg102 and not Lys101; note that the hydrogen is pointing straight at the carbonyl oxygen of Arg102. Thus, in the final turn of this α-helix, the hydrogen-bonding pattern has changed from i to $i + 4$ to i to $i + 3$, more characteristic of the 3_{10} turn. This change in the hydrogen-bonding pattern is frequently found in the C-terminal caps of α-helices.

Despite the hydrogen bonding of some atoms of the peptide bonds in the N- and C-terminal caps, there are still some non-hydrogen-bonded atoms. This will confer a small positive charge at the N-terminus of the α-helix and a negative one at the C-terminus. These charges are referred to as the dipole moment of an α-helix; thus, the dipole moment measures the extent of charge separation in a protein. The dipole moment is strengthened by the hydrogen bonding of the peptide bonds of the α-helix and by the environment of the helix (Hol et al. 1978; Hol 1985).

The presence of the dipole in α-helices is also of appreciable functional significance. It has been known for almost 40 years that the positive dipoles of the N-termini of α-helices can bind to phosphate groups, for example, as part of the recognition of co-enzymes such as NADH (Hol et al. 1978). We will see in ▶ Chap. 6 (▶ Sect. 6.4) that this is also the case for $p21^{ras}$ when we look at the interaction with its bound ligand GDP (see Exercise 5.12.7). In addition, in papain, the positive dipole on the central α-helix affects the properties of Cys25, the active site nucleophile, by lowering the pK_a from the usual value of 8 for cysteine residues to around 4 (Hol et al. 1978). This enables the enzyme to be active over a much broader pH range than would normally be the case (Menard et al. 1990).

5.7.3 Arrangements of the Side-Chains in α-Helices

While looking at the properties of the helices on the "protein" page of the PDBsum entry for 1q21, you will most probably have noticed the helical wheel drawings above the tables. However, there are also helical wheel drawings in the pdf file underneath the table. I find that these are more informative. The helical wheels show the positions of the amino acid side-chain relative to the helical axis and thus whether there are any groupings of amino acids on a particular face of a helix. The α-helix from residues Ser127-Tyr137 in $p21^{ras}$ is an excellent example. All of the seven polar and charged amino acids are on one face of the α-helix whereas the four apolar residues are on the opposite face. An α-helix with the amino acids arranged in this way is termed an amphipathic helix. Looking at the position of this helix in $p21^{ras}$, one sees the polar face exposed to the solvent whereas the apolar one faces inward and interacts with the side-chains of residues from the β-sheet 6. Amphipathic helices can often tether proteins to membranes, with the apolar face binding to the membrane and the polar face interacting with the solvent (Karanasios et al. 2010). When bound in this way, a protein may be able to sense or even modulate the curvature of a membrane (Drin et al. 2007).

5.7.4 3_{10} Helices

We looked above at the 3_{10} helical turn. Two or more consecutive 3_{10} turns constitute a 3_{10} helix. Vieira-Pires and Morais-Cabral (2010) have reviewed and listed proteins in which longer 3_{10} helices occur. One of these 3_{10} helices, formed by residues 150–158 of the rather obscure enzyme dienelactone hydrolase, is shown in ◘ Fig. 5.14; you should now be able to draw these diagrams. The 3_{10} helix of dienelactone hydrolase has two turns with three amino acids, each amino acid turning 120° (◘ Fig. 5.14a). As a consequence, the positions of the backbone atoms are very straightforward to visualize. If we look down the 3_{10} helix as in ◘ Fig. 5.14b, we see, for instance, that each C_α atom has another directly below it.

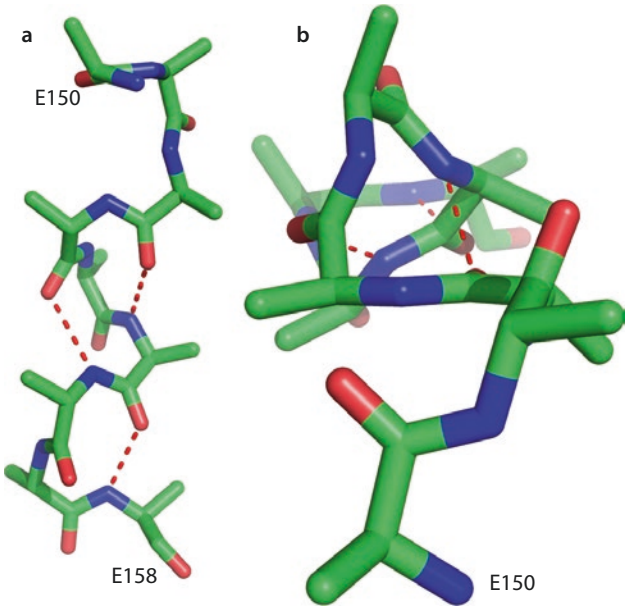


Fig. 5.14 Two views of the 3_{10} helix of dienelactone hydrolase (1din) formed by residues 150–158. **a** Side view. **b** Top view (rotated 90° from A about the *x*-axis) from the N-terminus

Hence, the side-chain of the residues will also be directly below each other so that the side-chains lack the stagger found in the α-helix. This arrangement of the side-chains in the 3_{10} helix is thought to be less optimal than in the α-helix. It may also be one reason why 3_{10} helices are less stable. Further reasons for its instability are that the polar contacts are not as straight as in the α-helix and that the 3_{10} helix is thinner than the α-helix. Nevertheless, 3_{10} helices between 6 residues and 11 residues in length have been observed in protein structures (Enkhbayar et al. 2006). However, these observed helices were often irregular.

Despite their irregularity, 3_{10} helices can play important functional roles. For instance, a 3_{10} helix in certain voltage-gated membrane proteins that transport potassium ions is an integral part of the voltage sensing mechanism. Read the above-mentioned review to find out more (Vieira-Pires and Morais-Cabral 2010).

5.7.5 Pi Helices

A pi (π) helix has a hydrogen-bonding pattern from the carbonyl oxygen of residue *i* to the amide hydrogen of residue *i* + 5. Such helices do occur in proteins and can be observed in Fig. 2 of the article by Gonzalez et al. (2012).

5.8 Conformational Preferences of Amino Acids

Many studies have been carried out to determine the conformational preferences of the amino acids. Table 5.6 summarizes these values determined from high-resolution structures for α-helices, β-sheets, and reverse turns. The amino acids that have a preference for an α-helical conformation are generally polar, with the exception of leucine, methionine,

Table 5.6 Conformational preferences for the 20 biogenic amino acids

Residue	α-helix (P_α)	β-sheet (P_β)	Reverse turn (P_t)
Glu	**1.59**	0.52	1.01
Ala	**1.41**	0.72	0.82
Leu	**1.34**	1.22	0.57
Met	**1.30**	1.14	0.52
Gln	**1.27**	0.98	0.84
Lys	**1.23**	0.69	1.07
Arg	**1.21**	0.84	0.97
His	**1.05**	0.80	0.81
Val	0.90	**1.87**	0.41
Ile	1.09	**1.67**	0.47
Tyr	0.74	**1.45**	0.76
Cys	0.66	**1.40**	0.54
Trp	1.02	**1.35**	0.65
Phe	1.16	**1.33**	0.59
Thr	0.76	**1.17**	0.90
Gly	0.43	0.58	**1.77**
Asn	0.76	0.48	**1.34**
Pro	0.34	0.31	**1.32**
Ser	0.57	0.96	**1.22**
Asp	0.99	0.39	**1.24**

Data taken from (Creighton 2013)

and phenylalanine. The polar residues are expected to be staggered away from each other as they are on the outside of the helix. But why are leucine and methionine found more often in α-helices? This preference reflects the presence of leucine in α-helical coil structures termed leucine zippers that we will explore in ▶ Chap. 7. In these structures, leucine residues form an interface between two polypeptide chains, thus finding themselves in a hydrophobic environment.

In contrast, residues that prefer a β-sheet conformation are mostly hydrophobic. The explanation lies in the proximity of the side-chains to each other in the β-sheet. We can examine this using ◘ Fig. 3.13 of ▶ Chap. 3. This figure shows the details of the side-chains in the β-sheet of papain. To determine the proximity of the side-chains in

the β-sheet, return to this drawing and locate residue Ile173. Use the selection command "sele i. 173 around 4" in PyMOL to find the six hydrophobic residues within a circle of 4 Å around Ile173. The proximity of the side-chains stems from two factors. The first is that, seen from the edge of the β-sheet, the side-chains of every second residue in one β-strand point in the same direction. Secondly, the side-chains in adjacent β-strands simultaneously all point upward or downward. This proximity favors the presence of hydrophobic residues because only a steric fit is required to pack them together. If a large number of hydrophilic residues were to be present, it would be necessary to arrange their charged and polar side-chains to avoid electrostatic clashes. This would be much more difficult and the resulting structure much less stable. Indeed, the packing of hydrophobic side-chains is an important factor for the stability of the β-sheet.

The frequency for each amino acid in each conformation was calculated from the frequency of the particular amino acid in a conformation divided by the fraction of all residues in that conformation. Thus, a value over 1.00 indicates that an amino acid prefers a particular conformation; the converse is true for values less than 1.00. My students find this concept difficult to grasp. It is perhaps easier to understand when the probability is expressed in mathematical terms:

$$P = \frac{f\left(\text{amino acid } x \text{ in a particular conformation}\right)}{f\left(\text{all residues in a particular conformation}\right)}$$

As an illustration, let us calculate the P_α value of 1.59 for glutamate for an imaginary protein. Suppose that in this protein 40% of the amino acids are in an α-helical conformation. In order for the P_α to be 1.59, 64% of the glutamate residues must be in an α-helical conformation. The equation is for glutamate is then

$$P_\alpha \text{ for glutamate} = \frac{f\left(\text{glutamate in } \alpha-\text{helices}\right)}{f\left(\text{all residues in } \alpha-\text{helices}\right)} = \frac{0.64}{0.4} = 1.59.$$

Take-Home Messages

The α-helix and the β-sheet are the two most important secondary structure elements in proteins. Both elements can vary in their torsion angles so that they may deviate considerably from theoretically optimal values. Further, the uncertainty in defining the positions of the hydrogen bonds leads to uncertainty in the assignment of the secondary structure of a protein. Currently available algorithms may offer different assignments which should be examined in the structure of the protein. Other secondary structures such as the 3_{10} helix and the β-bridge are also observed. Secondary structure elements are connected by turns which can be characterized by their hydrogen-bonding properties and torsional angles. Often several turns may follow each other, resulting in multiple turns. In addition to glycine and proline, small polar residues such as asparagine, aspartic acid, and serine are often present in turns because of their abilities to form hydrogen bonds and adopt unusual torsional angles.

5.9 Summary of PDB Information Discussed in This Chapter

The records "HELIX" and "SHEET" indicate the residues that form the secondary structure elements. For helices, there is information on the length and type of helix present. For sheets, there is information on the numbers on residues in the β-strands and whether the β-strands are parallel or anti-parallel. In addition, the records list atoms that are hydrogen bonded between the strands. The residues involved in secondary structure elements as well as turns can be visualized in the PDB using the output from the DSSP algorithm. The DSSP output can be reached by clicking on the button "View Sequence & DSSP Image" in the section "Chain Downloadable Files" in the sequence tab.

5.10 Summary of PyMOL Commands Introduced in This Chapter

5.10.1 General and Settings

To see all parameters that can be changed in PyMOL, go to "Setting" in the drop-down menu and then "Edit All." Most never need to be changed, but you may prefer to use this menu to change the settings mentioned in this book. Underneath "Edit All" is "Colors" which list all the colors in PyMOL and allows you to view and edit them.

5.10.2 Organization

Set dash_length

5.10.3 Viewing

h_add

Examples

h_add, sele (uses an algorithm to add hydrogens if not in PDB file. This is a difficult task, so not always correct. By the way, why is it a difficult task?)
show sticks, sele (makes the hydrogens visible)

5.10.4 Selecting

Examples

select 9pap and (i. 5–7 i. 164–167 i. 170–175 i. 185–191 i. 148–149)
create papain_sheet_1, sele (creates new object just with selected parts of original)

5.10.5 Coloring

Examples

util.cbac sele (color by atom, carbon atom blue)
color magenta, ss s (colors only β-strands)
color cyan, ss l+ (colors only loops)

5.10.6 Changing

Examples

alter 9pap and i. 148–149, ss='S' (to force PyMOL to assign secondary structure)
rebuild

5.11 Further Reading

5.11.1 Books

Hager T (1998) Linus Pauling and the chemistry of life. Oxford portraits in science. Oxford University Press, New York

Hall KT (2014) The man in the monkeynut coat. William Astbury and the forgotten road to the double-helix. Oxford University Press, Oxford

Perutz MF (1998) I wish I'd made you angrier earlier: essays on science, scientists and humanity. Cold Spring Harbor Press, Cold Spring Harbor

Kyle J (2007) Structure in protein chemistry. Chapter 6: Atomic details. Garland Science, New York

5.11.2 Online Resources

5.11.2.1 PDBsum

► www.ebi.ac.uk/pdbsum

In this chapter, we used the pages of "protein tab" of PDBsum entries to examine the positions, residues, and variations in secondary structure elements and classify the types of turns found in protein structures.

5.11.2.2 General

Run DSSP on-line yourself at:

► http://www.cmbi.ru.nl/xssp/

A more detailed description of the secondary structure assignments in the PDB:

► https://www.cgl.ucsf.edu/chimera/docs/UsersGuide/tutorials/pdbintro.html

5.12 Exercises (Table 5.7)

Table 5.7 Structures for the exercises in this chapter

PDB entry identifier	PDB entry title	Organism	Reference
7rsa	Structure of phosphate-free ribonuclease A at 1.26 Angstroms	*Bos taurus*	Wlodawer et al. (1988)
9pap	Structure of papain refined at 1.65 Angstroms resolution	*Carica papaya*	Kamphuis et al. (1984)
1din	Dienelactone hydrolase at 2.8 Angstroms	*Pseudomonas knackmussii*	Pathak and Ollis (1990)
1mcp	Phosphocholine binding immunoglobulin binding Fab MC/PC603. An X-ray diffraction study at 2.7 Angstroms	*Mus musculus*	Satow et al. (1986)
1q21	Crystal structures at 2.2 Angstrom resolution of the catalytic domains of normal ras protein and an oncogenic mutant complexed with GSP	*Homo sapiens*	Tong et al. (1991)

5

5.12.1 We looked at RNase A at the start of the chapter but then continued with papain and $p21^{ras}$. Examine the structure and turns of RNase A using DSSP, PDB, and PDBsum. If you just draw the backbone, can you recognize the secondary structure elements just from their patterns? Can you find any amphipathic helices?

5.12.2 Staying with RNase A, if you look at the torsion angles in the VADAR algorithm, can you identify the secondary structure elements from them? You can also take a peek at VADAR algorithm to see how it assigns secondary structure (► http://vadar.wishartlab.com/help/info.struct.txt).

5.12.3 Draw the structures of RNase A and papain as cartoons and color according to secondary structure. Use the information in Figs. 5.3 and 5.4 to color in red the residues that hydrogen bond between the strands.

5.12.4 We saw above the irregularity in the 3_{10} helix in the structure 1din. Have a look at α-helices in this molecule (e.g., show the α-helices as sticks and look at the polar contacts) and compare them to the 3_{10} helix. Investigate their properties from PDBsum. Look at the ends for any unusual variation or the presence of N- and C-terminal caps. Use the VADAR website (► http://vadar.wishartlab.com/) to examine the torsion angles, especially at the N- and C-termini of the α-helices.

5.12.5 Among the first molecules to have their structures determined was that of an Fab fragment of an immunoglobulin molecule (i.e., the part that

interacts with an antigen) (Padlan et al. 1973; Poljak 1975). Both the heavy and light chains are composed of β-sheets. To intensify your understanding of secondary structure elements, have a look at the structure (PDB identifier 1mcp) from the papers cited above. Make a cartoon and color the molecule according to chain (util.cbc) to visualize the β-sheets and the loops that hold them together. Then, color each chain with a spectrum from the N- to the C-terminus (util.rainbow chain A). What do you notice about the loops at the N-terminus of the chains? Hide the cartoon and show the backbone as sticks. Examine the arrangement of the strands. Are they parallel or anti-parallel? Look at the DSSP assignments and the information in PDBsum. What sort of turns can you find between the strands? Are the strands consecutive in sequence or not? We will look more closely at this structure in ► Chap. 8 (► Sect. 8.1).

? 5.12.6 I mentioned in ► Sect. 2.1 that the "Molecule of the Month" feature at the RCSB PDB entry site. $p21^{ras}$ was one once "Molecule of the Month"; you can examine the article at ► http://pdb101.rcsb.org/motm/148 to prepare yourself for ► Chap. 6 which examines the interaction of this protein with its ligands and binding partners in more detail.

? 5.12.7 Perhaps you think we have squeezed everything out of $p21^{ras}$. No, it still has plenty to offer. To link ► Chaps. 5 and 6, visualize the ligand GDP in the structure 1q21. You can easily do this in PyMOL with the command "show sticks, i. 180" (GDP is marked as residue 180 in the PDF file). The phosphate groups are in orange; which structure do you think is binding and neutralizing their negative charge? Use the command "show sticks, br. i. 180 around 4" to find the answers. The argument "br." means by residue.

? 5.12.8 Explain the difference between a dipole moment and a dielectric constant.

References

Astbury WT, Woods HJ (1930) The X-ray interpretation of the structure and elastic properties of hair keratin. Nature 126:913–914. https://doi.org/10.1038/126913b0

Astbury WT, Woods HJ (1934) X-ray studies of the structure of hair, wool, and related fabrics II – the molecular structure and elastic properties of hair keratin. Philos Trans R Soc Lond 232:333–U367. https://doi.org/10.1098/rsta.1934.0010

Chou PY, Fasman GD (1977) Beta-turns in proteins. J Mol Biol 115(2):135–175

Creighton TE (2013) Proteins: structure and molecular properties, 2nd edn. W H Freeman & Co, San Francisco

Drin G, Casella JF, Gautier R, Boehmer T, Schwartz TU, Antonny B (2007) A general amphipathic alpha-helical motif for sensing membrane curvature. Nat Struct Mol Biol 14(2):138–146. https://doi.org/10.1038/nsmb1194

Enkhbayar P, Hikichi K, Osaki M, Kretsinger RH, Matsushima N (2006) 3(10)-helices in proteins are parahelices. Proteins 64(3):691–699. https://doi.org/10.1002/prot.21026

Fersht AR, Shi JP, Knill-Jones J, Lowe DM, Wilkinson AJ, Blow DM, Brick P, Carter P, Waye MM, Winter G (1985) Hydrogen bonding and biological specificity analysed by protein engineering. Nature 314(6008):235–238

Gonzalez A, Cordomi A, Caltabiano G, Pardo L (2012) Impact of helix irregularities on sequence alignment and homology modeling of G protein-coupled receptors. Chembiochem 13(10):1393–1399. https://doi.org/10.1002/cbic.201200189
Hager T (1998) Linus Pauling and the chemistry of life. Oxford portraits in science. Oxford University Press, New York
Hol WG (1985) Effects of the alpha-helix dipole upon the functioning and structure of proteins and peptides. Adv Biophys 19:133–165
Hol WG, van Duijnen PT, Berendsen HJ (1978) The alpha-helix dipole and the properties of proteins. Nature 273(5662):443–446
Huggins ML (1943) The structure of fibrous proteins. Chem Rev 32:195–218. https://doi.org/10.1021/cr60102a002
Hutchinson EG, Thornton JM (1994) A revised set of potentials for beta-turn formation in proteins. Protein Sci 3(12):2207–2216. https://doi.org/10.1002/pro.5560031206
Hutchinson EG, Thornton JM (1996) PROMOTIF--a program to identify and analyze structural motifs in proteins. Protein Sci 5(2):212–220. https://doi.org/10.1002/pro.5560050204
Kabsch W, Sander C (1983a) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22(12):2577–2637. https://doi.org/10.1002/bip.360221211
Kabsch W, Sander C (1983b) How good are predictions of protein secondary structure? FEBS Lett 155(2):179–182
Kamphuis IG, Kalk KH, Swarte MB, Drenth J (1984) Structure of papain refined at 1.65 A resolution. J Mol Biol 179(2):233–256
Karanasios E, Han GS, Xu Z, Carman GM, Siniossoglou S (2010) A phosphorylation-regulated amphipathic helix controls the membrane translocation and function of the yeast phosphatidate phosphatase. Proc Natl Acad Sci U S A 107(41):17539–17544. https://doi.org/10.1073/pnas.1007974107
Martin J, Letellier G, Marin A, Taly JF, de Brevern AG, Gibrat JF (2005) Protein secondary structure assignment revisited: a detailed analysis of different assignment methods. BMC Struct Biol 5:17. https://doi.org/10.1186/1472-6807-5-17
Menard R, Khouri HE, Plouffe C, Dupras R, Ripoll D, Vernet T, Tessier DC, Laliberte F, Thomas DY, Storer AC (1990) A protein engineering study of the role of aspartate 158 in the catalytic mechanism of papain. Biochemistry 29(28):6706–6713
Milner-White EJ, Ross BM, Ismail R, Belhadj-Mostefa K, Poet R, (1988) One type of gamma-turn, rather than the other gives rise to chain-reversal in proteins. Journal of Molecular Biology 204 (3):777–782
Pace CN, Shirley BA, McNutt M, Gajiwala K (1996) Forces contributing to the conformational stability of proteins. FASEB J 10(1):75–83
Padlan EA, Segal DM, Spande TF, Davies DR, Rudikoff S, Potter M (1973) Structure at 4.5 A resolution of a phosphorylcholine-binding fab. Nat New Biol 245(145):165–167
Panasik N Jr, Fleming PJ, Rose GD (2005) Hydrogen-bonded turns in proteins: the case for a recount. Protein Sci 14(11):2910–2914. https://doi.org/10.1110/ps.051625305
Pathak D, Ollis D (1990) Refined structure of dienelactone hydrolase at 1.8 A. J Mol Biol 214(2):497–525
Pauling L (1940) A theory of the structure and process of formation of antibodies. J Am Chem Soc 62:2643–2657. https://doi.org/10.1021/ja01867a018
Pauling L, Corey RB (1951) The pleated sheet, a new layer configuration of polypeptide chains. Proc Natl Acad Sci U S A 37(5):251–256
Pauling L, Corey RB, Branson HR (1951) The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. Proc Natl Acad Sci U S A 37(4):205–211
Pavone V, Gaeta G, Lombardi A, Nastri F, Maglio O, Isernia C, Saviano M (1996) Discovering protein secondary structures: classification and description of isolated alpha-turns. Biopolymers 38(6):705–721. https://doi.org/10.1002/(SICI)1097-0282(199606)38:6<705::AID-BIP3>3.0.CO;2-V
Perutz MF (1951) New X-ray evidence on the configuration of polypeptide chains. Nature 167(4261):1053–1054
Poljak RJ (1975) Three-dimensional structure, function and genetic control of immunoglobulins. Nature 256(5516):373–376
Richardson JS, Richardson DC (1988) Amino acid preferences for specific locations at the ends of alpha helices. Science 240(4859):1648–1652
Rose GD, Gierasch LM, Smith JA (1985) Turns in peptides and proteins. Adv Protein Chem 37:1–109

Satow Y, Cohen GH, Padlan EA, Davies DR (1986) Phosphocholine binding immunoglobulin Fab McPC603. An X-ray diffraction study at 2.7 A. J Mol Biol 190(4):593–604

Tong LA, de Vos AM, Milburn MV, Kim SH (1991) Crystal structures at 2.2 Å resolution of the catalytic domains of normal ras protein and an oncogenic mutant complexed with GDP. J Mol Biol 217(3): 503–516

Vieira-Pires RS, Morais-Cabral JH (2010) 3(10) helices in channels and other membrane proteins. J Gen Physiol 136(6):585–592. https://doi.org/10.1085/jgp.201010508

Weatherford DW, Salemme FR (1979) Conformations of twisted parallel beta-sheets and the origin of chirality in protein structures. Proc Natl Acad Sci USA 76(1):19–23

Wilmot CM, Thornton JM (1988) Analysis and prediction of the different types of beta-turn in proteins. J Mol Biol 203(1):221–232

Wilmot CM, Thornton JM (1990) Beta-turns and their distortions: a proposed new nomenclature. Protein Eng 3(6):479–493

Wlodawer A, Svensson LA, Sjolin L, Gilliland GL (1988) Structure of phosphate-free ribonuclease A refined at 1.26 A. Biochemistry 27(8):2705–2717