Secondary Structure in Properdin of the Complement Cascade and Related Proteins: A Study by Fourier Transform Infrared Spectroscopy†

Stephen J. Perkins,*,‡,§ Adam S. Nealis,‡,§ Parvez I. Haris,§ Dennis Chapman,§ Dimitrios Goundis,∥ and Kenneth B. M. Reid∥

Department of Biochemistry and Chemistry and Department of Protein and Molecular Biology, Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2PF, U.K., and MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

Received February 15, 1989; Revised Manuscript Received May 12, 1989

ABSTRACT: Six structural repeat motifs of 58 amino acids are found in the sequence of both mouse and human properdins. Twelve more examples of the motif are available from the sequences of thrombospondin, the terminal complement components, and the thrombospondin-related anonymous protein. The averaged Robson and Chou–Fasman secondary structure predictions show that there are 57–66% turn and 19–38% β-sheet structures in the typical repeat motif. The high amount of turn structure is consistent with Gly, Pro, Cys, and Ser being the four most abundant amino acid residues in properdin. Comparisons with sequences found in the circumsporozoite protein from several species of malaria parasites show that their sequences and secondary structures strongly coincide only in a 18-residue segment. Further secondary structure analysis utilized Fourier transform infrared spectroscopy of human properdin in ^2H_2O buffers. These show a broad amide I band that, after second-derivative and deconvolution calculations, is shown to be composed of several components. Two at 1633 and 1683 cm^{-1} are strong evidence for β-sheet structure, although overlap from β-turns can also contribute. The presence of β-turn structure is indicated by absorptions at 1662–1675 and 1645 cm^{-1}. The properdin structure contains substantial quantities of β-sheet and β-turn structures, which is consistent with the secondary structure predictions and amino acid compositions. The length of the repeat motif is estimated as 3.3–4.3 nm, and an estimated 14–22% of nonexchanged amide protons reside in properdin. This is suggestive of a high degree of solvent accessibility in the structure.

Properdin is a member of the alternative pathway of complement activation (Reid, 1986). It has a polydisperse structural form and is found as a mixture of cyclic "head-to-tail" dimers, trimers, and tetramers of a 53 600 M_r monomer (Reid, 1981; DiScipio, 1982; Smith et al., 1984). Its physiological role is to stabilize the C3/C5 convertase produced by the alternative pathway, which is the complex formed between C3b and Bb. Sequence studies show that properdin is constructed from six structurally similar motifs of 58 residues each (Goundis & Reid, 1988; Goundis, 1988), together with an additional 39–42 residues at the N- and C-terminii. Since these motifs are homologous to three structural domains originally found in thrombospondin (Lawler & Hynes, 1986), they are termed "thrombospondin repeats" (TSRs). These TSR domains are also found (DiScipio et al., 1988; Patthy, 1988) in two copies in each of the complement components C7, C8α, and C8β and one copy in C9, all of which are components of the membrane attack complex of complement. Related sequences are also found in the circumsporozoite protein from several species of malaria parasites.

The elucidation of the secondary structure of the TSR is thus of importance for several groups of medically important proteins. Prediction methods based on primary sequences conventionally result in accuracies in the range of 50–60% (Busetta & Hospital, 1982; Kabsch & Sander, 1983; Nishikawa, 1983; Perkins et al., 1989; Argos, 1989). Improvements in their accuracy to 65–75% can be obtained by averaging techniques (Argos, 1989), either by the use of templates to recognize supersecondary structures (Taylor & Thornton, 1983, 1984) or by the use of a number of aligned, homologous sequences (Zvelebil et al., 1987; Sawyer et al., 1988; Perkins et al., 1988).

Fourier transform infrared (FT-IR) spectroscopy is able to report on the secondary structures of proteins (Susi & Byler, 1986; Lee & Chapman, 1986; Surewicz & Mantsch, 1988). The structural information is obtained from the infrared bands of the conformation-sensitive amide bands, particularly the amide I band. Studies on polypeptides and proteins have shown that there is a correlation between the amide I frequency and the type of secondary structure present. Another well-established technique for obtaining information about protein secondary structure is circular dichroism (CD) spectroscopy. FT-IR spectroscopy, however, has several advantages over CD spectroscopy. The FT-IR method is not affected by the presence of high amounts of disulfide bridges in the protein, there is no problem in defining a base line in the FT-IR spectrum, and different parts of the FT-IR spectrum can be clearly resolved. Unlike CD, the FT-IR method does not suffer from light-scattering problems, and it can be also used to obtain information about the amide proton exchange rates within proteins. These advantages were demonstrated in an earlier study on human factor H of complement. The CD spectrum (DiScipio & Hugli, 1982) was unable to identify a β-sheet structure that was, however, readily apparent by FT-IR spectroscopy (Perkins et al., 1988).

Structural analyses on the TSR are most ideally facilitated by the use of human properdin. Approximately 80% of its

† S.J.P. is a Wellcome Trust Lecturer in Biochemistry. A.S.N. is supported by the Wellcome Trust. P.I.H. is supported by the Science and Engineering Research Council through a CASE award sponsored by Smith, Kline & French Research Ltd.

‡ Department of Biochemistry and Chemistry, Royal Free Hospital School of Medicine.

§ Department of Protein and Molecular Biology, Royal Free Hospital School of Medicine.

∥ University of Oxford.

0006-2960/89/0428-7176$01.50/0 © 1989 American Chemical Society

structure is comprised of TSR motifs, and these will be dominant in the spectroscopic investigation. As for factor H, no secondary structure could be inferred from CD studies (Smith et al., 1984). The features of the CD spectra were not those characteristic of α-helix, β-sheet, or random coil structure. The full mouse sequence (Goundis & Reid, 1988), however, shows that there are six Cys residues per TSR and four more in each of the N- and C-terminal regions, and Cys is in fact the third most abundant amino acid in properdin. It is possible that this high disulfide content has obscured the interpretation of the CD spectra. The joint use of secondary structure averaging and FT-IR is a promising and useful approach. In this study it is used to identify an unusual secondary structure in the TSR. Further comparisons are made with an 18-residue homologous polypeptide sequence that is found in the circumsporozoite protein of malaria parasites.

Materials and Methods

(*a*) *Sequence Alignments.* Thirty-one sequences for the thrombospondin repeat were used in this study. These are given as follows in the order in which they appear in Figure 1: (a) human thrombospondin (Lawler & Hynes, 1986), mouse properdin (Goundis & Reid, 1988), human properdin (Reid & Gagnon, 1981; Goundis, 1988), human C7 (DiScipio et al., 1988), human C8α (Rao et al., 1987), human C8β (Haefliger et al., 1987; Howard et al., 1987), human C9 (DiScipio et al., 1984; Stanley et al., 1985), and thrombospondin-related anonymous protein (TRAP) from *Plasmodium falciparum* (Robson et al., 1988); (b) circumsporozoite protein of *Plasmodium knowlesi* (Ozaki et al., 1983), *P. falciparum* (Dame et al., 1984), *Plasmodium yoelii* (Lal et al., 1987), *Plasmodium vivax* (Arnot et al., 1985), *Plasmodium cynomolgi* (Galinski et al., 1987), *Plasmodium berghei* (Eichinger et al., 1986), and *Plasmodium brasilianum* (Lal et al., 1988). Human properdin has been sequenced at the gene level from the N-terminal end as in Figure 1 up to residues HCPG in the second TSR, and the remaining sequence shown (Figures 1 and 4) is a partial amino acid sequence at the protein level (Goundis, 1988). The TSR sequences for human properdin are thus 82% complete. Secondary structures were predicted by the Chou and Fasman (1978) method and by the Robson method (Garnier et al., 1978) and were averaged (Perkins et al., 1988, 1989). The Robson calculation is biased toward β-structure by reason of the infrared spectroscopy results to be described below.

(*b*) *Purification of Human Properdin.* A Sepharose anti-properdin monoclonal antibody column was prepared by coupling a mouse monoclonal antibody (supplied by Dr. C. Koch, Statens Seruminstitut, Denmark) to CNBr-activated Sepharose 4B. The column (1 cm × 5 cm), containing 5 mg of antibody/mL of Sepharose 4B, was equilibrated in phosphate buffer saline, pH 7.4 (PBS). Serum (2 L) that had been extensively dialyzed against PBS was preabsorbed with Sepharose 4B and then applied to the monoclonal antibody column which was then extensively washed with PBS. Any bound material was eluted first with PBS made 3 M with respect to NaCl and then with 0.2 M glycine–0.5 M NaCl, pH 2.5. The pH 2.5 eluate was immediately neutralized by the addition of 1 M NaOH. This fraction was reapplied to the affinity column and eluted by using the same conditions as before. Final purification of properdin was achieved by gel filtration on Sephacryl S-300, as detailed below. The overall yield was approximately 50% of that expected with 10 mg of properdin being obtained from 1 L of serum.

When properdin was applied to a Sephacryl S-300 column (85 × 2.5 cm), the protein eluted in three distinct peaks between 160 and 235 mL. This is a reproducible observation, and it reflects the fact that in human serum properdin exists as dimers, trimers, tetramers, and higher oligomers (trimers being the most prevalent) (Reid, 1981; DiScipio, 1982, Smith et al., 1984; Pangburn, 1989). No significant amount of the monomeric form was seen on gel filtration in nondissociating conditions. Protein solutions were concentrated at 4 °C by ultrafiltration under compressed air using Amicon Diaflo ultrafiltration cells and PM10 membranes (Amicon BV, Oosterhout, Holland). It was found that to concentrate the oligomeric forms of properdin after S-300 gel filtration to above 0.5 mg/mL, 0.2 M glycine had to be added to the sample buffer (150 mM NaCl, 12 mM sodium phosphate, 0.2 mM EDTA, pH 7.4).

(*c*) *FT-IR Spectroscopy.* Infrared spectra were recorded with a Perkin-Elmer 1750 FT-IR spectrometer equipped with a TGS detector. A Perkin-Elmer Model 7300 data station was used for data acquisition, storage, and analysis. Samples were dialyzed at 4 °C into 500 mM NaCl and 12 mM sodium phosphate, pH 7.0, in 100% ^2H_2O for 36–48 h with four changes of buffer. Samples were placed in a thermostated Beckmann FH-01 CFT microcell fitted with CaF_2 windows and a 100-μm Teflon spacer. The protein samples and their buffers were measured with identical scanning parameters after equilibration at 20 °C for 15 min. Spectra were obtained by signal averaging of 2500 scans at a resolution of 4 cm^{-1} (6 h 20 min). The ^2H_2O buffer subtraction was carefully carried out such that a straight base line around 2000–1800 cm^{-1} was obtained. Further details about spectral subtraction are given in Haris et al. (1986) and Mitchell et al. (1988). To reduce noise, the second derivative (rate of change of slope) was calculated over a 13 data point range (13 cm^{-1}), as for factor H of complement (Perkins et al., 1988). Spectral deconvolution was performed by using the Perkin-Elmer ENHANCE function, which is analogous to the method developed by Kauppinen et al. (1981). The deconvolution parameters used for the amide I band were $\sigma = 9$ (half-width at half-height) and $K = 2.5$.

Human properdin and human immunoglobulin IgG (Sigma I4506) were prepared in KBr disks after extensive dialysis at 4° C into 200 mM KBr and lyophilization. About 0.6 mg of properdin was ground into 250 mg of dried KBr and transferred to a die. The mixture was subjected to a pressure of 10 tons under vacuum for 10 min to form a disk, and this disk was transferred to a circular specimen holder for FT-IR studies. Likewise, 2.0–2.6 mg of IgG was used. Second-derivative analysis was carried out as described above for the samples in ^2H_2O. Due to the much greater bandwidth of the properdin–KBr spectrum, σ was set as 11 for deconvolution. The much improved signal to noise ratio enabled a K value of 3.0 to be applied.

Results and Discussion

(*a*) *Alignment of the TSR Sequences.* The TSR occurs not only in thrombospondin and in mouse and human properdin but also in one or two copies in 12 other proteins that are involved either in the circumsporozoite or blood stages of the malaria parasite or in the terminal complement (Goundis & Reid, 1988; Robson et al., 1988). A total of 31 sequences from 15 proteins are compared in Figure 1. Their alignment was chosen to minimize gaps and maximize homologies, and this scheme was employed for the averaging of the TSR secondary structure predictions. In thrombospondin and properdin, the maximum length of the TSR was determined as 71 residues. Since 58 of these 71 positions are significantly occupied in the sequences under investigation, this was set as the standard

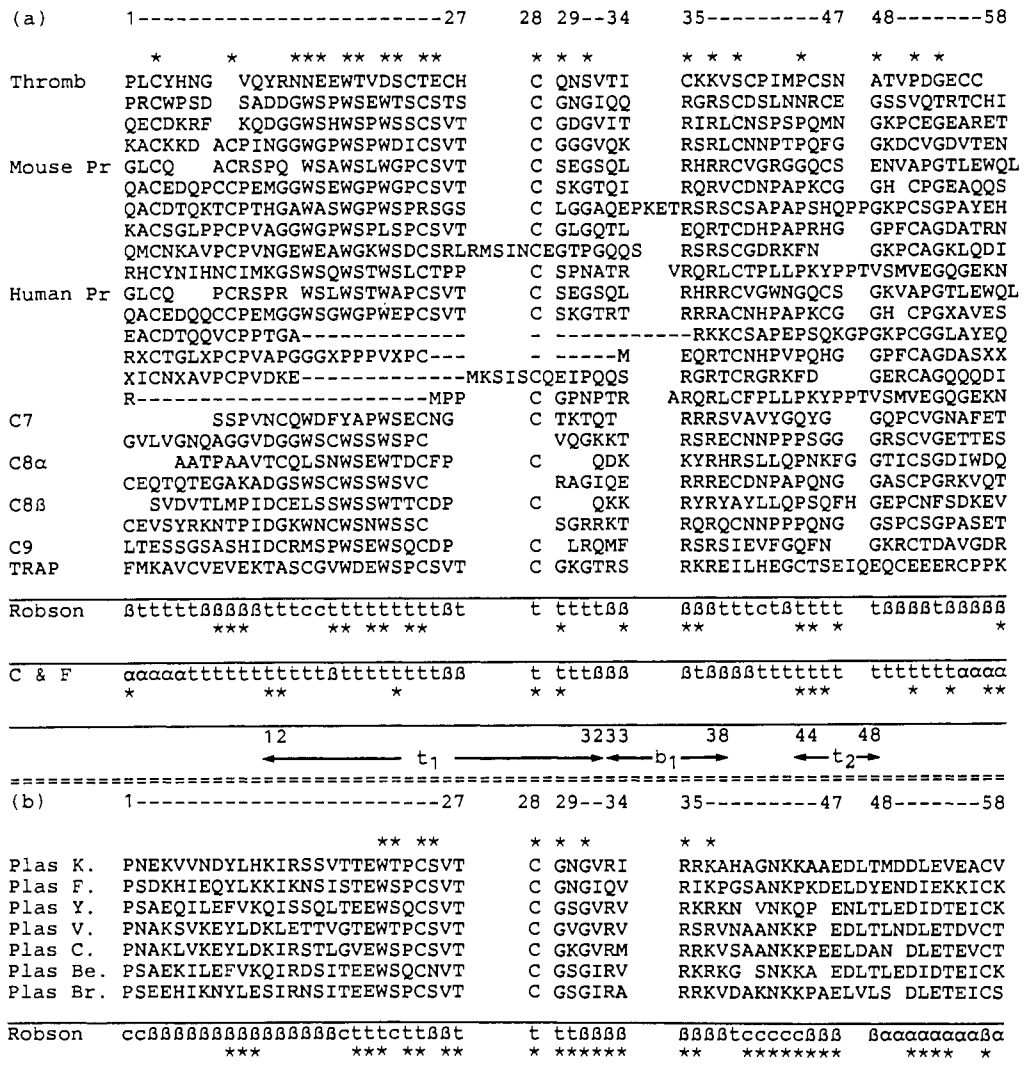

FIGURE 1: Alignment of the 31 sequences for the thrombospondin repeat. The source of the sequences is given under Materials and Methods. The alignment is divided into two parts, (a) and (b), and residue groups that are conserved in 60% or more of the sequences in either part are asterisked above the sequences. The residue groups are as defined and employed previously (Perkins et al., 1988, 1989). The secondary structure assignment is given below each part, and this is asterisked if this is maintained in 75% or more of the sequences. Symbols: α, α-helix; β, β-strand; t, turn; c, coil. In (a), the sequence of human properdin is incomplete; residues that are missing by homology with the mouse sequence are denoted by dashes and unknown residues by X.

length of a TSR, and Figure 1 is numbered accordingly.

The 31 sequences were divided into two groups. A good homology of the 7 circumsporozoite sequences in Figure 1b with the 24 sequences in Figure 1a extends over only residues 20–37. The two groups were accordingly analyzed separately. The TRAP sequence shows a slightly improved homology with the thrombospondin and properdin sequences and so is analyzed with them. Three TSRs were identified in the original sequencing study of thrombospondin, together with parts of a fourth TSR (Lawler & Hynes, 1986). The realignment shown in Figure 1a shows that in fact a complete fourth TSR of weaker homology but with conserved Cys and Trp residues at positions 3 and 18 can be identified in thrombospondin, and this is included in the TSR analyses. The alignment of Figure 1 is similar to that proposed in Lawler and Hynes (1986), Goundis and Reid (1988), Robson et al. (1988), and Patthy (1988). However, the presumed location of the TSR N-terminus is not clear; this is 8–14 residues earlier in Figure 1 relative to that proposed in other studies.

Residue conservation is summarized in Figure 1a. Of the 58 residues, 21 are conserved to 60% or better in the total of up to 31 sequences that are available. The most conserved ones are Trp-18, Cys-24, Cys-28, and Arg-37 (90% or better). A second well-conserved set (80–90%) corresponds to Ser-16, Ser-19, Trp-21, Arg-35, Gly-48, and Gly-53. Of the remaining 11 conserved positions, 4 of these correspond to Cys-3, Cys-9, Cys-39, and Cys-51. On average there are six Cys residues per TSR in properdin and thrombospondin, and four Cys residues per TSR in the terminal complement components. However, none of these can be rigorously said to occur in identical positions in all 24 sequences. A further four conserved residues correspond to the subset of the smallest residues (Gly, Ala, Ser). In this context it is noteworthy that each of Trp-15, Trp-18, and Trp-21 is very frequently followed by a Ser residue or a related one.

Averaged secondary structures (Perkins et al., 1988) were predicted by both the Robson and the Chou–Fasman methods to serve as a mutual control of their outcome. Figure 1 shows that both predictions lead to a preponderance of turns in two regions t_1 and t_2 in the TSRs and one β-sheet region b_1. For the 24 TSR sequences, the Robson method predicts 0% α-helix, 38% β-sheet, 57% turn, and 5% coil. The Chou–Fasman method predicts 16% α-helix, 19% β-sheet, and 66% turn. The largest divergences between the two predictions occur at the N- and C-terminal ends of the TSR domains. Possible explanations are that (a) the structure is less well defined at the

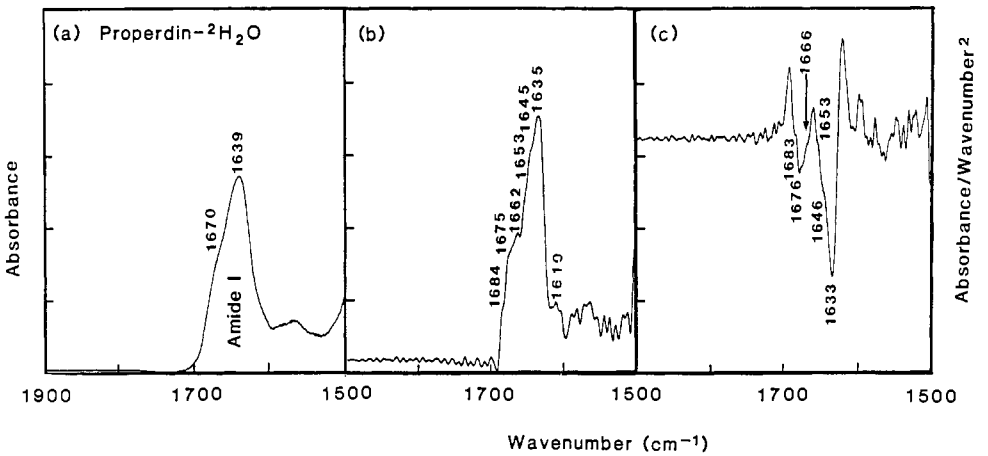

FIGURE 2: FT-IR spectra of human complement component properdin in the amide I region. Panels a–c correspond to properdin in ^2H_2O buffer. Sample concentrations ranged between 0.7 and 1.2 mg/mL. (a) Difference spectrum of deuterated tetrameric properdin (P_4) in 0.5 M NaCl and 12 mM sodium phosphate buffer in ^2H_2O at p^{2}H 7.0, 20 °C. (b) Deconvoluted spectrum and (c) second-derivative spectrum of that shown in (a).

start and end of a domain in a multidomain structure or (b) residue conservation is weaker in this region. As expected, the prediction for the 7 circumsporozoite sequences is in good agreement with that for the TSRs only between the positions 20 and 37 noted above (Figure 1b). Since this region may mimic the action of the complement components in question, to evade the immune system of the host, the implication is that residues 20–37 may constitute the active site of the TSR domain.

(*b*) *Infrared Spectroscopy of Properdin.* FT-IR spectroscopy was performed to investigate the secondary structure further. Such FT-IR studies normally require protein concentrations of 10–30 mg/mL (e.g., the complement component factor H; Perkins et al., 1988). Properdin was, however, only available in concentrations up to 1.5 mg/mL because of its low solubility. Data collection in H_2O buffers was not possible due to the strong intensity of the H_2O band. Data could only be obtained in ^2H_2O buffers. There, signal to noise ratios were further improved by use of a 100-μm sample thickness in place of 6 μm and by signal averaging over 2500 scans (6 h 20 min) rather than 400 scans [cf. Perkins et al. (1988)]. Figure 2 shows that satisfactory measurements were possible, despite a noisy background in parts b and c.

The FT-IR spectrum of deuterated properdin in its tetrameric form (P_4) in ^2H_2O buffer is displayed in Figure 2a. The amide I maximum is at 1639 cm^{-1}, and a prominent shoulder is visible near 1670 cm^{-1}. Second-derivative analysis of the absorbance spectrum reveals only two main components (Figure 2c). The major amide I component is at 1633 cm^{-1}, with a minor component clearly seen at 1676 cm^{-1}. Minor components are seen at 1683, 1666–1663, 1653, and 1646 cm^{-1}. The spectrum was also analyzed by deconvolution, and this gives similar results (Figure 2b).

Properdin is found in three forms, which are dimers (P_2), trimers (P_3), and tetramers (P_4). These are reproducibly obtained from gel filtration based on Sephacryl S-300 column, when the protein is eluted in three distinct peaks between 160 and 235 mL (Reid, 1981; DiScipio, 1982; Smith et al., 1984). All three peaks give an identical pattern on SDS-PAGE under both nonreducing and reducing conditions. The trimer is the most prevalent form. The FT-IR absorbance spectra of these three forms were recorded in ^2H_2O buffers to show that all three forms are very similar. As a result of the low concentrations and the consequential poor signal to noise ratio, small differences in their secondary structures would not have been detected. It is concluded that the secondary structures of these forms are similar and are not perturbed by the noncovalent interactions between the subunits.

The decrease in intensity of the FT-IR amide II band in ^2H_2O buffers can be used as a measure of ^1H–^2H exchange in peptide groups. The amide II band is located between 1600 and 1500 cm^{-1}. For properdin, the intensity of the band in this region is very weak (Figure 2a), and this suggests that most of the amide protons have exchanged. Since the spectrum in H_2O was not measurable, the amide II/amide I intensity ratio for undeuterated properdin could not be directly obtained for calculation of the extent of amide proton exchange. However, this ratio is generally accepted as 0.45 for undeuterated soluble proteins (Osborne & Nabedryk-Viala, 1982). By use of this value and assuming that the amide II band in properdin is located at 1550 cm^{-1}, the upper limit of nonexchanged amide protons is estimated as 14–22%. This is similar to that of 23% for factor H (Perkins et al., 1988) and is consistent with a high degree of solvent accessibility.

(*c*) *FT-IR Band Assignments in Properdin.* The spectral assignments of the properdin FT-IR spectrum in ^2H_2O were primarily based on reports in Krimm and Bandekar (1980, 1986), Susi and Byler (1986), Lee and Chapman (1986), Olinger et al. (1986), Haris et al. (1986), Yasui et al. (1986), Haris and Chapman (1988), and Surewicz and Mantsch (1988). In ^2H_2O, the position of the main amide I component at 1633 cm^{-1} is consistent with the presence of β-sheet. However, β-turn structures can also absorb in this region. The amide I component centered near 1633–1635 cm^{-1} is quite broad (Figure 2b,c) in comparison to the corresponding band observed in immunoglobulin IgG (Figure 3b), which is a β-sheet protein. It is likely that absorbances from both β-sheet and β-turn structures are present in properdin and overlap to such a degree that more detailed band structures cannot be observed after second-derivative or deconvolution analysis. No estimates of the proportion of these two structures in properdin are thus possible.

The second major band at 1676 cm^{-1} in ^2H_2O probably arises from β-turn structure. The presence of several absorbances near 1660-1683 cm^{-1} suggests the existence of β-turn structure, in particular those near 1663–1666 cm^{-1}. The intensity of absorbance near 1670 cm^{-1} is greater for properdin compared to other proteins known to have a high proportion of β-sheet. This is exemplified by immunoglobulin IgG in Figure 3a,b [see also Wasacz et al. (1987)], where stronger absorptions are seen around 1680 cm^{-1} and at higher frequencies. Absorbances near 1670–1675 cm^{-1} have been as-

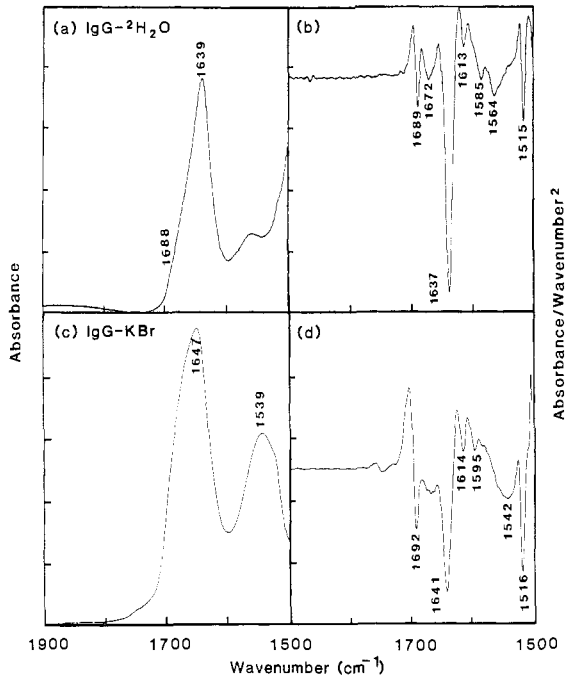

FIGURE 3: FT-IR spectra of human immunoglobulin IgG. (a) The difference spectrum of deuterated IgG is measured at a concentration of 20 mg/mL by using a 6-μm tin spacer, with 400 scans averaged at a resolution of 4 cm^{-1}. The buffer is phosphate buffer saline in ^2H_2O, pH 7.4. (b) Second-derivative spectrum of that in (a), calculated in the same way as for properdin. (c) The difference spectrum of protonated IgG in a KBr disk is shown. (d) Second-derivative spectrum of that in (c).

signed to turn structures, and those at higher frequencies to β-sheet structures. Thus, the weak features near 1683 cm^{-1} are best assigned to the higher frequency component of the antiparallel β-sheet (Figure 2c vs Figure 3b). Again, an estimate of the ratio of turn or β-sheet structure is not possible because of band overlap.

Of the remaining bands in Figure 2c, the weak feature at 1653 cm^{-1} is compatible with a small amount of α-helix. Another weak component near 1644–1645 cm^{-1} is difficult to assign precisely, but may arise from either random coil or β-turns.

As a further control of the assignments, protonated trimeric properdin was studied in KBr disks. While improved signal to noise ratios are obtained, the amide I band is broader than in ^2H_2O, and properdin activity after lyophilization (as measured by a complement alternative pathway hemolytic assay) was reduced to 8% of that of a control sample which had been dialyzed into 200 mM KBr but not lyophilized. Despite this loss of activity, bands could be observed after the deconvolution and second-derivative analyzes that support the analysis of the ^2H_2O spectrum. Components were seen at 1633–1634 and 1690 cm^{-1}, which represent the low- and high-frequency components of an antiparallel β-sheet, respectively, and the presence of turn structure is indicated by a prominent band near 1664–1666 cm^{-1}. IgG was also measured in KBr disks for comparison with the ^2H_2O spectrum (Figure 3c,d). As for properdin, the amide I band in KBr is broader compared to that in ^2H_2O, and its maximum shifts by 8 cm^{-1} to a lower frequency in ^2H_2O. This shift is smaller than that for properdin. The second-derivative analyses of IgG in ^2H_2O and in the KBr disk show similar bands (part b vs part d of Figure 3), indicating that the antiparallel β-sheet structure is retained in the solid state.

To summarize, the band positions observed in ^2H_2O suggest that properdin contains substantial amounts of turn and β-sheet

```
                    1                                          42
Mouse Pr  N-term    DVPLCFTQYEESSGRCKGLLGRDIRVEDCCLNAAYAFQEHDG
Human Pr            DPVLCFTQYEESSGKCKGLLGGGVSVEDCCLNTAFAYQKRSG

Mouse               ßßßßßßßß3ttttttttßßßßßßßßßßßß3tttaa3ßßßßttt
Human               ßßßßßßßß3ttttttttßßßßttcßßßßßßßßßßßßßßttt

                  400                                        440
Mouse Pr  C-term    VTFWGTPRPLCEALQGQKLVVEEKRSCLHVPVCKDPEEKKP
Human Pr            VTFSGRPLPRCEELQGQKLVVEEKRPCLHVPACKDPEEEE

Mouse               ßßttcccccttßßßatßßßßßßttttßßßßßßcctttct
Human               ßßßttcccctßßßattaßßßßttttta3ßßßttcccaaaa


Mouse Pr:     Ile  6; Phe  5; Val 19; Leu 24; Trp 19; Met  5; Ala 26;
              Gly 47; Cys 43; Tyr  5; Pro 46.
              Thr 21; Ser 39; His 11; Glu 25; Asn 12; Gln 28; Asp 13;
              Lys 20; Arg 27.
```

FIGURE 4: Sequences for the N-terminal and C-terminal segments of properdin. These correspond to the remaining sequence of properdin not shown in Figure 1a. Below these, the β-biased Robson predictions are shown and total 5% α-helix, 55% β-strand, 30% turn, and 10% coil. Symbols are defined in Figure 1. The amino acid composition of mouse properdin is shown, where the residues are listed in order from the most hydrophobic to the most hydrophilic. On the assumption of two triantennary complex type oligosaccharides (28 residues), the total M_r for a properdin subunit is 53 600 (10.7% CHO), and the predicted absorption coefficient (1%, 280 nm, 1 cm) is 22.8, following the procedure of Perkins (1986).

structure. A major contribution from α-helical and random coil structures is not evident. This is fully consistent with the secondary structure predictions.

(*d*) *Conclusions.* The joint results from secondary structure predictions, the amino acid compositions, and FT-IR spectroscopy are evidence that the predominant conformation in the TSR is a turn structure, together with β-structure. The 57–66% content of turns predicted in the secondary structure of the TSR is high for a globular protein, since this is usually about 25%. However, such a content is not unknown for globular proteins; X-ray crystallography, proton NMR, and vibrational and circular dichroism spectroscopy have indicated that metallothioneins have extensive turn structures (Pande et al., 1986; Kägi & Schäffer, 1988). In properdin, these turns can be correlated with the large number of conserved Ser, Cys, Pro, and Gly residues (Figure 1), since β-turns are sterically demanding on amino acid side chains. The amino acid composition (Figure 4) shows that Ser and Gly are two of the four most abundant residues. This can be correlated with the prospensity of Ser and Gly residues to occur in type I β-turns and Gly in type II β-turns (Wilmot & Thornton, 1988). Likewise, Pro residues are favored in β-turns, and Pro is the second most abundant amino acid in properdin. It is, however, not possible to predict the precise location of β-turns in the consensus TSR sequence on the basis of the residue patterns noted for type I and II turns by Wilmot and Thornton (1988).

Some information on the gross three-dimensional structure of the TSR is available from combining the sequence data with electron microscopy. Properdin has been visualized as an extended globular-like structure by electron microscopy (Smith et al., 1984). This suggested that a properdin protomer is 26 nm in length. If it is assumed that this length corresponds to a linear arrangement of six TSR domains, then each TSR is 4.3 nm long. If, alternatively, this length corresponds to six TSRs and the N- and C-terminal domains (all eight of approximately equal sizes), each TSR is 3.25 nm long. These dimensions are compatible with lengths of between 16 and 29 nm that are estimated for a thin flexible strand connecting the two large globular domains of thrombospondin by electron microscopy (Lawler et al., 1985; Galvin et al., 1985). It is assumed that this strand corresponds to a linear arrangement of four TSR domains and three epidermal growth factor like domains. If each TSR is 3.25–4.3 nm long and if each epidermal growth factor domain is 2–3.5 nm long (Perkins & Nealis, 1989), the total length of the strand is predicted to be

in the region 19–29 nm, in good agreement with that observed.

The TSR structure can be compared with the secondary structure predictions and the TSR volume. The possible β-strand b_1 between positions 33 and 38 is 2.1 nm long, since the α-carbon atoms are separated by 0.35 nm. There is thus sufficient space for the β-strand b_1 since the TSR long axis is of length 3.25–4.3 nm. The dry volume of the TSR (Chothia, 1975; Perkins, 1986) is calculated as 66.1 nm^3 from the sequence (Figure 4). If circular in shape, the mean diameter of the TSR will be 1.8 nm. This is less than the value of 2.5–3.0 nm estimated from electron microscopy. This difference may suggest that the properdin protomer has an elongated cross section. In turn, this suggests that for its size properdin has a large surface area exposed to solvent. This is consistent with the high degree of ^1H–^2H exchange seen by FT-IR spectroscopy. In conclusion, properdin accords well with the general observation that many of the complement components have unusually extended protein structures.

Registry No. Pr, 11016-39-0; C7, 80295-57-4; C8α, 80295-58-5; C9, 80295-59-6.

References

Argos, P. (1989) in *Protein structure: a practical approach* (Creighton, T. E., Ed.) pp 169–189, IRL Press, Oxford, U.K.

Arnot, D. E., Barnwell, J. W., Tam, J. P., Nussenzweig, V., Nussenzweig, R. S., & Enea, V. (1985) *Science 230*, 815–818.

Busetta, B., & Hospital, M. (1982) *Biochim. Biophys. Acta 701*, 111–118.

Chothia, C. (1975) *Nature (London) 254*, 304–308.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45–148.

Dame, J. B., Williams, J. L., McCutchan, T. F., Weber, J. L., Wirtz, R. A., Hockmeyer, W. T., Maloy, W. L., Haynes, J. D., Schneider, I., Roberts, D., Sanders, G. S., Reddy, E. P., Diggs, C. L., & Miller, L. H. (1984) *Science 225*, 593–599.

DiScipio, R. G. (1982) *Mol. Immunol. 19*, 631–635.

DiScipio, R. G., & Hugli, T. E. (1982) *Biochim. Biophys. Acta 709*, 58–64.

DiScipio, R. G., Gehring, M. R., Podack E. R., Kan, C. C., Hugli, T. E., & Fey, G. H. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7298–7302.

DiScipio, R. G., Chakravarti, D. N., Müller-Eberhard, H. J., & Fey, G. H. (1988) *J. Biol. Chem. 263*, 549–560.

Eichinger, D. J., Arnot, D. E., Tam, J. P., Nussenzweig, V., & Enea, V. (1986) *Mol. Cell. Biol. 6*, 3965–3972.

Galinski, M. R., Arnot, D. E., Cochrane, A. H., Barnwell, J. W., Nussenzweig, R. S., & Enea, V. (1987) *Cell 48*, 311–319.

Galvin, N. J., Dixit, V. M., O'Rourke, K. M., Santoro, S. A., Grant, G. A., & Frazier, W. A. (1985) *J. Cell Biol. 101*, 1434–1441.

Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol. 120*, 97–120.

Goundis, D. (1988) D.Phil. Thesis, University of Oxford.

Goundis, D., & Reid, K. B. M. (1988) *Nature (London) 335*, 82–85.

Haefliger, J. A., Tschopp, J., Nardelli, D., Wahli, W., Kocher, H. P., Tosi, M., & Stanley, K. K. (1987) *Biochemistry 26*, 3551–3556.

Haris, P. I., & Chapman, D. (1988) *Biochim. Biophys. Acta 943*, 375–380.

Haris, P. I., Lee, D. C., & Chapman, D. (1986) *Biochim. Biophys. Acta 874*, 255–265.

Howard, O. M. Z., Rao, A. G., & Sodetz, J. M (1987) *Biochemistry 26*, 3565–3570.

Kabsch, W., & Sander, C. (1983) *FEBS Lett. 155*, 179–182.

Kägi, J. H. R., & Schäffer, A. (1988) *Biochemistry 27*, 8509–8515.

Kauppinen, J. K., Moffatt, D. J., Mantsch, H. H., & Cameron, D. G. (1981) *Appl. Spectrosc. 35*, 271–276.

Krimm, S., & Bandekar, J. (1980) *Biopolymers 19*, 1–29.

Krimm, S., & Bandekar, J. (1986) *Adv. Protein Chem. 38*, 181–364.

Lal, A. A., de la Cruz, V. F., Welsh, J. A., Charoenvit, Y., Maloy, W. L., & McCutchan, T. F. (1987) *J. Biol. Chem. 262*, 2937–2940.

Lal, A. A., de la Cruz, V. F., Collins, W. E., Campbell, G. H., Procell, P. M., & McCutchan, T. F. (1988) *J. Biol. Chem. 263*, 5495–5498.

Lawler, J., & Hynes, R. O. (1986) *J. Cell Biol. 103*, 1635–1648.

Lawler, J., Derick, L. H., Connolly, J. E., Chen, J. H., & Chao, F. C. (1985) *J. Biol. Chem. 260*, 3762–3772.

Lee, D. C., & Chapman, D. (1986) *Biosci. Rep. 6*, 235–256.

Mitchell, R. C., Haris, P. I., Fallowfield, C., Kelling, D. J., & Chapman, D. (1988) *Biochim. Biophys. Acta 941*, 31–38.

Nishikawa, K. (1983) *Biochim. Biophys. Acta 748*, 285–299.

Olinger, J. M., Hill, D. M., Jakobsen, R. J., & Brody, R. S. (1986) *Biochim. Biophys. Acta 869*, 89–98.

Osborne, N. B., & Nabedryk-Viala, E. (1982) *Methods Enzymol. 88*, 676–680.

Ozaki, L. S., Svec, P., Nussenzweig, R. S., Nussenzweig, V., & Godson, G. N. (1983) *Cell 34*, 815–822.

Pande, J., Pande, C., Gilg, D., Vasak, M., Callender, R., & Kägi, J. H. R. (1986) *Biochemistry 25*, 5526–5532.

Pangburn, M. K. (1989) *J. Immunol. 142*, 202–207.

Patthy, L. (1988) *J. Mol. Biol. 202*, 689–696.

Perkins, S. J. (1986) *Eur. J. Biochem. 157*, 169–180.

Perkins, S. J., & Nealis, A. S. (1989) *Biochem. J.* (in press).

Perkins, S. J., Haris, P. I., Sim, R. B., & Chapman, D. (1988) *Biochemistry 27*, 4004–4012.

Perkins, S. J., Nealis, A. S., Dudhia, J., & Hardingham, T. E. (1989) *J. Mol. Biol. 206*, 737–754.

Rao, A. G., Howard, O. M. Z., Ng, S. C., Whitehead, A. S., Colten, H. R., & Sodetz, J. M. (1987) *Biochemistry 26*, 3556–3564.

Reid, K. B. M. (1981) *Methods Enzymol. 80*, 143–150.

Reid, K. B. M. (1986) *Essays Biochem. 27*, 27–68.

Reid, K. B. M., & Gagnon, J. (1981) *Mol. Immunol. 18*, 949–959.

Robson, K. J. H., Hall, J. R. S., Jennings, M. W., Harris, T. J. R., Marsh, K., Newbold, C. I., Tate, V. E., & Weatherall, D. J. (1988) *Nature (London) 335*, 79–82.

Sawyer, L., Fothergill-Gilmore, L. A., & Freemont, P. S. (1988) *Biochem. J. 249*, 789–793.

Smith, C. A., Pangburn, M. K., Vogel, C. W., & Müller-Eberhard, H. J. (1984) *J. Biol. Chem. 259*, 4582–4588.

Stanley, K. K., Kocher, H. P., Luzio, J. P., Jackson, P., & Tschopp, J. (1985) *EMBO J. 4*, 375–382.

Surewicz, W. K., & Mantsch, H. H. (1988) *Biochim. Biophys. Acta 952*, 115–130.

Susi, H., & Byler, D. M. (1986) *Methods Enzymol. 130*, 290–311.

Taylor, W. R., & Thornton, J. M. (1983) *Nature (London) 301*, 540–542.

Taylor, W. R., & Thornton, J. M. (1984) *J. Mol. Biol. 173*, 487–514.

Wasacz, F. M., Olinger, J. M., & Jakobsen, R. J. (1987) *Biochemistry 26*, 1464–1470.
Wilmot, C. M., & Thornton, J. M. (1988) *J. Mol. Biol. 203*, 221–232.
Yasui, S. C., Keiderling, T. A., Bonora, G. M., & Toniolo, C. (1986) *Biopolymers 25*, 79–89.
Zvelebilk, M. J., Barton, G. J., Taylor, W. R., & Sternberg, M. J. E. (1987) *J. Mol. Biol. 195*, 957–961.

Conformational and Biological Properties of the Ala[10] Analogue of Human Des-Trp[1],Nle[12]-minigastrin†

Stefano Mammi,‡ Maria Teresa Foffani,‡ Evaristo Peggion,*,‡ Jean Claude Galleyrand,§ Jean Pierre Bali,§ Mario Simonetti,∥,⊥ Walter Göhring,∥ Luis Moroder,∥ and Erich Wünsch∥

Biopolymer Research Center, Department of Organic Chemistry, University of Padova, Via Marzolo 1, 35131 Padova, Italy, Centre de Recherche de Biochemie Macromoleculaire CNRS, LP-8402 INSERM U-249, Faculte de Pharmacie, 15 Avenue C. Flahault, 34060 Montpellier Cedex, France, and Max-Planck-Institut für Biochemie, Abteilung Peptidchemie, 8033 Martinsried bei München, Federal Republic of Germany

Received December 13, 1988; Revised Manuscript Received April 10, 1989

ABSTRACT: Synthesis, conformation, and biological properties of the Ala[10] analogue of des-Trp[1],Nle[12]-minigastrin are reported. Replacement of the Gly residue in the original sequence with Ala remarkably changes the conformational preference of the hormone in trifluoroethanol. CD and NMR results indicate that the conformational change is mainly located in the C-terminal portion of the molecule, with probable extension of the N-terminal α-helix throughout the entire sequence. The structural modification causes a 10-fold decrease in the biological potency of the hormone, which is about as active as the C-terminal tetrapeptide amide. These findings support our previous hypothesis that the optimal bioactive conformation of the native hormone is U-shaped, with mutual interactions among the two end segments.

Previous investigations on the synthetic analogues of human gastrin, des-Trp[1],Nle[12]-minigastrin

```
     2   3-7   8   9   10  11  13  13  14
H-Leu-[Glu]5-Ala-Tyr-Gly-Trp-Nle-Asp-Phe-NH2
```

and Nle[15]-little gastrin

```
 1   2   3   4   5   6-10  11  12  13  14  15  16  17
pGlu-Gly-Pro-Trp-Leu-[Glu]5-Ala-Tyr-Gly-Trp-Nle-Asp-Phe-NH2
```

led to the hypothesis that the conformation assumed by these hormones in TFE[1] (Peggion et al., 1985; Peggion & Foffani, 1985) or in aqueous solutions containing detergent micelles (Mammi et al., 1987) is of biological relevance. A conformational model was proposed on the basis of CD results on both peptides and on their fragments, as well as high-resolution ^{1}H NMR studies on des-Trp[1],Nle[12]-minigastrin (Mammi et al., 1986, 1988). According to our hypothesis, the structure of this gastrin form is characterized by an α-helical segment at the N-terminus, comprising the –(Glu)$_5$– sequence, and by a bend in the central part of the molecule. The hydrogen bonds involving amide protons in the C–terminal region, suggested by NMR results, are compatible with the presence of a segment of 3_{10}-helix starting from Ala[8]. Thus, the proposed model involves two helical segments at the chain ends, stabilized by mutual interactions in a U-shaped structure.

In the proposed conformation, the Gly residue in position 10 appears to play a key role in determining the conformation of minigastrin. According to Chou and Fasman analysis (Chou & Fasman, 1977), replacement of Gly with a helix-forming residue would change completely the conformational preference of the peptide backbone, extending the α-helix throughout the entire molecule. This in turn should have a substantial effect on the biological potency of the hormone.

From previous structure–function correlation studies on gastrins, we have learned that replacement of the methionine residue by nearly isosteric norleucine is without any effect on the biological activity (Wünsch et al., 1982). Similarly, the hormonal potency was practically fully retained when the human minigastrin sequence was shortened by one residue at its N-terminus (Göhring et al., 1984). Correspondingly, for the present study a minigastrin analogue was synthesized in which the Gly residue at position 10 of our reference compound (des-Trp[1],Nle[12]-minigastrin) has been replaced by the helix-forming residue Ala. This gastrin analogue is used to test our working, structural hypothesis by comparative conformational and biological studies.

Experimental Procedures

Synthesis. Melting points were determined on a capillary melting point apparatus (Büchi) and are uncorrected. Optical rotations were measured in a thermostated 1-dm cell on a

†This work was supported by Consiglio Nazionale delle Ricerche, Italy.
‡University of Padova.
§Centre de Recherche de Biochemie Macromoleculaire CNRS.
∥Max-Planck-Institut für Biochemie.
⊥Present address: Biopolymer Research Center, Department of Organic Chemistry, University of Padova, Via Marzolo 1, 35131 Padova, Italy.

[1] Abbreviations: DCC, dicyclohexylcarbodiimide; TLC, thin-layer chromatography; HPLC, high-performance liquid chromatography; HPTLC, high-performance thin-layer chromatography; TFE, trifluoroethanol; DMF, dimethylformamide, NMP, *N*-methylpyrrolidone; NMR, nuclear magnetic resonance; CD, circular dichroism; UV, ultraviolet; TMS, tetramethylsilane; FID, free induction decay; TOCSY, total correlation spectroscopy; ROESY, rotating-frame nuclear Overhauser effect spectroscopy; CCK, cholecystokinin.

0006-2960/89/0428-7182$01.50/0 © 1989 American Chemical Society