Three-Dimensional Structure in Solution of Barwin, a Protein from Barley Seed†

Svend Ludvigsen‡ and Flemming M. Poulsen*

Carlsberg Laboratorium, Kemisk Afdeling, Gamle Carlsberg Vej 10, DK-2500 Valby, Copenhagen, Denmark

Received November 27, 1991; Revised Manuscript Received June 9, 1992

ABSTRACT: The solution structure of a 125-residue basic protein, barwin, has been determined using ^{1}H nuclear magnetic resonance spectroscopy. This protein is closely related to domains in proteins encoded by wound-induced genes in plants. Analysis of the ^{1}H nuclear Overhauser spectrum revealed the assignment of more than 1400 nuclear Overhauser effects. Twenty structures were calculated based on 676 nontrivial distance restraints, 152 torsion angle restraints (92 ϕ, 56 χ^1, and 4 ω for proline), and stereospecific assignments of 38 chiral centers, using distance geometry, simulated annealing, and restrained energy minimization. None of the distance restraints was violated by more than 0.5 Å in any of the 20 structures, and none of the torsion angle restraints was violated by more than 1° in any of the structures. The RMS difference between the calculated and target interproton distance restraints is 0.033 Å, and the average atomic RMS differences between the 20 structures and their geometric average are 1.23 Å for backbone atoms and 1.73 Å for all heavy atoms. The dominating structural feature of the protein is a well-defined four-stranded antiparallel β-sheet, two parallel β-sheets packed antiparallel to each other and four short α-helices. The binding site of barwin to the tetramer *N*-acetylglucosamine has been qualitatively investigated, and the dissociation constant of the complex has been determined using one-dimensional ^{1}H nuclear magnetic resonance spectroscopy.

In two preceding papers we have described the primary and secondary structure of barwin, a protein from barley seed, closely related to proteins encoded by wound-induced plant genes (Ludvigsen & Poulsen, 1992; Svensson et al., 1992). The purpose of the present study has been to determine the three-dimensional structure with the aim to add further background for the elucidation of the function of the protein. The sequence similarity between barwin and the C-terminal domain of the proteins encoded by the wound-induced genes in potato (Stanford et al., 1989) and rubber tree (Broekaert et al., 1990) is very close, suggesting that also the three-dimensional structures of this group of proteins are closely related and have the same structural fold. So the structure determination of barwin will likewise provide additional information about the structures of these. Furthermore, we have studied the binding site of the oligosaccharide β(1-4) tetramer of *N*-acetylglucosamine by the chemical shift pertubations of resonance lines in the ^{1}H NMR spectra. The tetramer is in fast exchange with the protein, and this property has allowed a determination of the dissociation constant for the complex.

The structure determination of barwin represents in itself a significant achievement. The protein consists of 125 amino acid residues, and as such it is one of the larger proteins for which an NMR structure determination has been applied; since the protein has not been cloned so far, isotope-enriched barwin has not been available for the present study. The structure determination, therefore, has relied on ^{1}H NMR spectroscopy only. Furthermore, since no crystal structure is yet available, the present NMR study represents the only source to the three-dimensional structure of this family of proteins that are most likely engaged in a plant defense mechanism.

METHODS

Restraints. A full description of the experimental procedures concerning the recording and processing of ^{1}H NMR spectra, assignments of these, measurements of coupling constants, stereospecific assignments, and integration of NOESY cross peaks were given in the preceding paper (Ludvigsen & Poulsen, 1992). The classification of distance restraints was based on well-established strategies (Williamson et al., 1985; Clore et al., 1985, 1986), by division into three classes: strong (1.8–2.7 Å), medium (1.8–3.3 Å), and weak (1.8–5.0 Å). For distance restraints involving methyl groups 0.5 Å was added to the upper limit. Average intensities of the sequential (i,i+1) H^{α}–H^{N} NOE in β-sheets were used as reference for strong peaks and sequential (i,i+1) H^{N}–H^{N} in α-helices for medium NOEs. NOEs which fully or partially overlapped in the spectra were usually assigned as weak if no reasonable judgment of the intensity classification could be made. Measurements of $^3J_{H^NH^\alpha}$ coupling constants (Ludvigsen et al., 1991; Ludvigsen & Poulsen, 1992) combined with the measurement of the intra H^{N}–H^{α} NOE were used to derive ϕ-angle restraints. For residues with a $^3J_{H^NH^\alpha}$ coupling constant greater than 7 Hz and with a weak intra H^{N}–H^{α} NOE the ϕ angles were restrained between –160° and –80°, for residues with a $^3J_{H^NH^\alpha}$ coupling constant lower than 6 Hz and a medium intra H^{N}–H^{α} NOE the ϕ-angle restraint was –90° to –10°, and for residues with $^3J_{H^NH^\alpha}$ coupling constants between 6.5 and 7.5 Hz together with a strong intra H^{N}–H^{α} NOE ϕ-angles were restrained between 30° and 90° (Kline et al., 1988; Ludvigsen et al., 1991). Side-chain torsion angles, χ^1, were restrained in one of the three staggered conformations 60°, 180°, or –60° with a range of ±60° as described previously.

Initially 1272 distance restraints (626 intra, 359 short, and 284 long range) and 152 torsion angles restrains (92 ϕ, 56 χ^1, and 4 ω for proline) could be determined. None of the distance restraints were trivial restraints like intra methylene, intra aromatic, and other similar restraints, but due to a noncomplete stereospecific assignment and generally many weak intra and

† The atomic coordinates of the structures have been deposited in the Brookhaven Protein Data Bank under Accession Numbers 1BW1, 1BW2, R1BW1MR, and R1BW2MR.

* Author to whom correspondence should be addressed.

‡ Present address: Novo Research Institute, Novo Nordisk, Novo Allé, DK-2880 Bagsværd, Copenhagen, Denmark.

0006-2960/92/0431-8783$03.00/0

short-range distance restraints, the total number of distance restraints containing useful structural information was lower. The distinction between meaningful distance restraints and others was carried out by the program DIANA (Günthert et al. 1991). After this procedure, the number of useful restraints used for the subsequent structure calculations was 676 distance restraints (133 intra, 278 short, and 265 long range) and all of the previously mentioned 152 torsion restraints. Eighteen hydrogen bond distance restraints (Table I) were used in the distance geometry calculation and the simulated annealing calculations but not in the final restrained minimization calculations.

Calculational Strategy. Structure calculations of barwin were performed using the hybrid method combining DISGEO (Crippen & Havel, 1975; Havel et al., 1983; Havel & Wüthrich, 1984, 1985; Sippl & Scheraga, 1986; Havel, 1986) and simulated annealing (Nilges et al., 1988) using a slightly modified cooling protocol and finally followed by restrained Powell minimization (Powell, 1977). The simulated annealing and the restrained Powell minimization were performed using the program X-PLOR version 2.1 (Brünger, 1988).

Using the DISGEO program, 300 structures were calculated to phase 2 of the DISGEO program, where each structure is represented by a subset of atoms. The distance restraints were corrected for pseudoatoms (Wüthrich et al., 1983). The 20 best converged structures were transferred to the X-PLOR program for further processing. An extended strand of barwin with all atoms was fitted to each of the DISGEO substructures. The simulated annealing protocol was the same as the one suggested by Nilges et al. (1988) except for the cooling protocol and the choice of distance and torsion restraint force constants. First, the 20 fitted structures were submitted to 200 cycles of unrestrained energy minimization. At this stage and in the rest of the simulated annealing calculations the potential functions used included terms of bonds (600 kcal mol^{-1} Å^{-2}), angles (500 kcal mol^{-1} radian^{-2}), impropers (500 kcal mol^{-1} radian^{-2}), planarity for the ω peptide bond (200 kcal mol^{-1} radian^{-2}), and a repulsive term using 0.8 times the normal van der Waals radius from the CHARMM potential (Brooks et al., 1983) with a variable force constant k_{vdW}. Additionally, the experimentally derived restraints were represented by a square well potential with variable force constants k_{NOE} and k_{CDIH} for the distance restraints and torsion angle restraints, respectively. Hydrogen bonds and the three cystine bridges were in the simulated annealing step treated as other distance restraints. For distance restraints including pseudoatoms a single potential averaged by $\langle r^{-6}\rangle^{1/6}$ was used. Fifty cycles of 0.075-ps Verlet dynamics (Verlet, 1967) were carried out at 1000 K. For each cycle k_{vdW} was increased by a factor of $400^{1/50}$ from 0.01 to 4.00 kcal mol^{-1} Å^{-2}, k_{NOE} increased by a factor of 2 from 0.5 to a maximum of 40.0 kcal mol^{-1} Å^{-2}, and k_{CDIH} increased by a factor of 2 from 2.5 to a maximum of 80.0 kcal mol^{-1} radian^{-2}. After this, the cooling process was performed by 28 cycles of 0.100 ps of molecular dynamics where the temperature was controlled by coupling to a heat bath. The temperature was lowered in each cycle by 25 K from 1000 to 300 K. Still with the same potential function, the structures were subjected to 200 cycles of restrained Powell energy minimization (Powell, 1977). After the simulated annealing protocol followed 800 cycles of restrained Powell energy minimization, now with the full CHARMM potential (Brooks et al., 1983) including potentials for bonds, angles, dihedrals, impropers, nonbonded Lennard-Jones interactions, hydrogen bonds, electrostatic interactions with a distance dependent dielectric constant (Gelin, 1976), and square well potentials for the experimentally derived restraints with force constants for k_{NOE} and k_{CDIH} at 40 kcal mol^{-1} Å^{-2} and 80 kcal mol^{-1} radian^{-2}, respectively. Cystine bonds were constructed and hydrogen bond restraints taken out prior to the 800 cycles of restrained energy minimization.

DISGEO calculations were performed on a micro VAX 3400 and X-PLOR calculations on a Silicon Graphics 320 power station or Stardent 1500 computers. Displays of structures were performed on the Stardent 1500 using a modified version

Table I: Hydrogen Bonds Found in the 20 Structures of Barwin[a]

donor backbone N	acceptor backbone O	secondary structure	d_{NO} (Å)	angle (N–H–O) (deg)
3[b,c]	118	β	3.0 ± 0.1	169 ± 4
5[b,c]	116	β		
6	116	β	2.9 ± 0.1	170 ± 5
8	114	β	3.0 ± 0.2	163 ± 8
18	15	T III	3.4 ± 0.5	139 ± 17
20	15	T	3.2 ± 0.4	133 ± 16
21	18	T III′	2.9 ± 0.3	148 ± 20
24	21	T I	3.2 ± 0.7	148 ± 20
27	24	T I	3.0 ± 0.8	158 ± 12
28[c]	24	T	3.6 ± 0.9	156 ± 6
32	29	T I	2.9 ± 0.4	148 ± 19
34[c]	31	3_{10}	2.9 ± 0.2	146 ± 8
35	31	α	2.7 ± 0.1	160 ± 12
36[c]	32	α	3.4 ± 0.5	137 ± 10
38	35	3_{10}	2.8 ± 0.3	147 ± 18
43	39	α	3.6 ± 0.3	161 ± 11
44[b]	40	α	2.8 ± 0.1	167 ± 5
45[b]	41	α	2.8 ± 0.1	149 ± 10
46[b]	42	α	3.0 ± 0.1	170 ± 5
49	81	β	2.9 ± 0.3	119 ± 23
51[b]	84	β	2.8 ± 0.1	155 ± 12
62[c]	59	T I	3.4 ± 0.7	138 ± 16
63	60	T I	3.7 ± 0.7	156 ± 21
65	62	T	3.0 ± 0.2	155 ± 13
66[b]	121	β	2.9 ± 0.1	147 ± 5
67[b]	80	β	2.8 ± 0.1	165 ± 9
68[b]	119	β	2.8 ± 0.1	168 ± 4
69[b]	78	β	2.9 ± 0.1	170 ± 5
70[b]	117	β	2.9 ± 0.1	173 ± 3
71[b]	76	β	2.8 ± 0.0	163 ± 6
74[c]	71	T III	3.1 ± 0.4	135 ± 16
75[c]	71	β	2.8 ± 0.1	155 ± 17
78[b]	69	β	2.8 ± 0.1	165 ± 6
80[b]	67	β	2.8 ± 0.1	169 ± 6
82	65	β	3.0 ± 0.1	158 ± 18
83[c]	49	β	3.0 ± 0.2	159 ± 14
84[b]	49	β	3.3 ± 0.3	163 ± 10
86	51	β	3.9 ± 0.9	140 ± 23
101	98	α	3.1 ± 0.3	141 ± 12
102	98	α	3.4 ± 0.3	149 ± 17
102	99	α	3.2 ± 0.4	135 ± 23
103	99	α	3.7 ± 0.6	133 ± 17
109[c]	106	3_{10}	3.6 ± 0.8	119 ± 24
110[c]	107	3_{10}	3.7 ± 0.6	137 ± 20
111	108	3_{10}	2.9 ± 0.3	134 ± 21
112[c]	108	T	3.0 ± 0.5	157 ± 20
114	8	β	3.0 ± 0.3	156 ± 14
116[b,c]	6	β	3.4 ± 0.3	164 ± 11
117[b]	70	β	3.0 ± 0.1	170 ± 5
118[b]	3	β	2.9 ± 0.1	164 ± 11
119[b]	68	β	2.9 ± 0.1	167 ± 5
120	1	β	3.4 ± 0.6	163 ± 12
121	66	β	2.8 ± 0.1	161 ± 4

[a] Average distances between N and O and average angles N–H–O are listed for each hydrogen bond with the standard deviation computed for the 20 structures. [b] Donors for which distance restraints have been used in the initial DISGEO and simulated annealing calculations. None of these were used in the final restrained minimization. [c] Fast amide proton exchange was observed. The notation for the secondary structure responsible for the hydrogen bonding is the following: α, α-helices; β, β-sheets; T, turns [turn type if identified uses the nomenclature of Venkatachalam (1968)]; 3_{10}, 3_{10}-helix identified on the basis of hydrogen bonding patterns.

Table II: Energies and Structural Deviations for 20 Structures of Barwin[a]

	energy term (number)									
	total	bond (1893)	angle (3375)	dihedral (865)	improper (666)	van der Waals	electrostatic	hydrogen bond	NOE restraints (676)	torsion restraints (153)
energy (kcal/mol)	−933 ± 72	65 ± 4	506 ± 25	594 ± 15	15.1 ± 1.8	−520 ± 17	−1501 ± 54	−121 ± 8	29 ± 8	0.48 ± 0.33
structural deviations		0.013 ± 0.0005 Å	3.17° ± 0.08°		0.38° ± 0.02°					

[a] Energies calculated with the CHARMM (Brooks et al., 1983) potential. Electrostatic potentials were calculated with a distance-dependent dielectric constant (Gelin, 1976). Force constants for distance restraints and torsion angle restraints are 40 kcal mol^{-1} Å^{-2} and 80 kcal mol^{-1} radian^{-2}, respectively. Numbers of bonds, angles, dihedrals, impropers, distance restraints, and torsion restraints are given in parentheses.

Table III: Violation of Distance Restraints as Average Number of Violations in Each Range for All 20 Structures[a]

range (Å)	intra $i = j$	short $1 \le \lvert i-j \rvert \le 5$	long $\lvert i-j \rvert > 5$	all
0.1–0.2	2.45	4.75	5.25	12.45
0.2–0.3	0.1	0.9	2.05	3.05
0.3–0.4	0	0.4	0.35	0.75
0.4–0.5	0	0.05	0.15	0.20
>0.5	0	0	0	0
RMS difference between calculated and target interproton distance restraints (Å)	0.026 ± 0.004	0.030 ± 0.007	0.037 ± 0.007	0.033 ± 0.005
no. of restraints	133	278	265	676

[a] Ranges are defined between residues i and j in the sequence.

of X-PLOR 1.2 (Kjær, unpublished data). All structural parameters for evaluation were calculated using X-PLOR except for some statistical analysis. The calculated average of the 20 final structures was only used as a computational reference. Average dihedral angles $\langle\theta\rangle$ were calculated from

$$\langle\theta\rangle = \mathrm{Im}\left(\ln\left(\frac{\sum_{k=1}^{N} e^{i\theta_k}}{\left|\sum_{k=1}^{N} e^{i\theta_k}\right|}\right)\right)$$

where N is the number of structures. Angular RMS values of θ are calculated by

$$\mathrm{rms}(\theta) = \cos^{-1}\left(\frac{1}{N}\left|\sum_{k=1}^{N} e^{i\theta_k}\right|\right)$$

Note that with this definition an angular RMS of 90° means that the angle is undefined.

Oligosaccharide Binding Studies. The binding of various saccharides [maltose, cellobiose, lactose, *N*-acetylglucosamine, and UDP–*N*-acetylglucosamine, and the β(1-4) tetramer of *N*-acetylglucosamine] has been studied by examination of chemical shift perturbation of ^{1}H resonance lines when these saccharides were added to the protein solution. In cases where perturbations were observed these were analyzed with respect to changes in line width and chemical shift values. The specific changes were then correlated with the resonance assignments to determine qualitatively the binding site. The protein concentration was kept constant at 0.41 mM and saccharides were added in increasing amounts from 0.4 to 8.0 ratios of ligand/protein.

RESULTS

The 20 structures were analyzed with respect to energies and deviations from ideal geometry (Table II), violation of distance restraints (Table III), atomic RMS differences, behavior of backbone dihedral angles, and hydrogen bonds (Table I). All atomic RMS differences in the following are average values for the difference between each of the 20 structures and the geometric average structure. The atomic RMS difference for all backbone atoms is 1.23 Å and for all heavy atoms 1.73 Å. All structures agreed well with the distance restraints (Table III); in fact, no distance restraint violation exceeded 0.5 Å in any of the 20 structures. No violation of torsion angle restraints greater than 1° was observed in any of the structures.

Description of the Structures. All 20 structures possessed the secondary structure elements predicted by the analysis of sequential NOEs, $^3J_{H^NH^\alpha}$ coupling constants, and hydrogen bonds (Ludvigsen & Poulsen, 1992). The major secondary structure component is the four-stranded antiparallel β-sheet (β-a) composed of residues 2–9 (β-Ia), 66–71 (β-IIIa), 75–81 (β-IVa), and 113–121 (β-IIa), arranged from one side as β-Ia, β-IIa, β-IIIa, and β-IVa. Four short strands formed a β-sheet (β-b) consisting of residues 10–13 (β-Ib), 48–52 (β-IIIb), 83–87 (β-IVb), and 90–94 (β-IIb), where β-Ib is packed parallel to β-IIb that again is antiparallel to β-IIIb which finally is parallel to β-IVb. The packing of this four-stranded β-sheet (β-Ib, β-IIb, β-IIIb, and β-IVb) was deduced from the C^α–C^α distance matrix (Figure 1); however, only the interaction between β-IIIb and β-IVb was observed in the NOESY spectra. Three small α-helices are present for residues 30–38 (α-I), 40–46 (α-II), and 97–103 (α-III) as expected from the secondary structure elucidation, but a fourth small α-helix is formed in several of the structures as well including residues 106–111 (α-IV). The hydrogen pattern of this helix resembles a 3_{10}-helix; however, some of these hydrogen bonds are quite weak. Helix α-I has as helix α-IV hydrogen bonds that indicate partly 3_{10}-helix character. Turns with hydrogen bonds (Table I) were recognized at sequence position 18–21 (type III′ turn), 21–24 (type I turn), 24–27 (type I turn), 60–63 (type I turn), 62–65 (type I turn), and 71–74 (type III turn). Two turns are further recognized on the basis of characteristic pairs of ϕ and ψ angles from 26 to 29 (type II turn) and from 122 to 125 (type II turn). Finally, a few other irregular turns exist deduced from the hydrogen bonds in Table I.

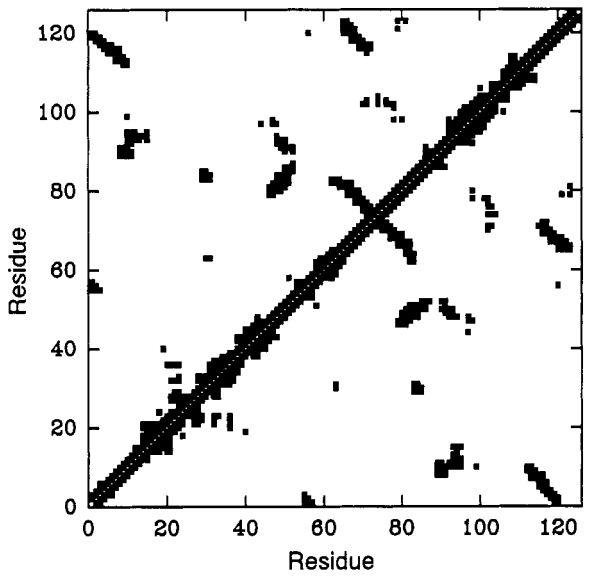

FIGURE 1: Distances shorter than 8 Å between C^{α} atoms in barwin shown in a matrix residue vs residue. Recognition of parallel and antiparallel β-sheet are easily traced from this representation.

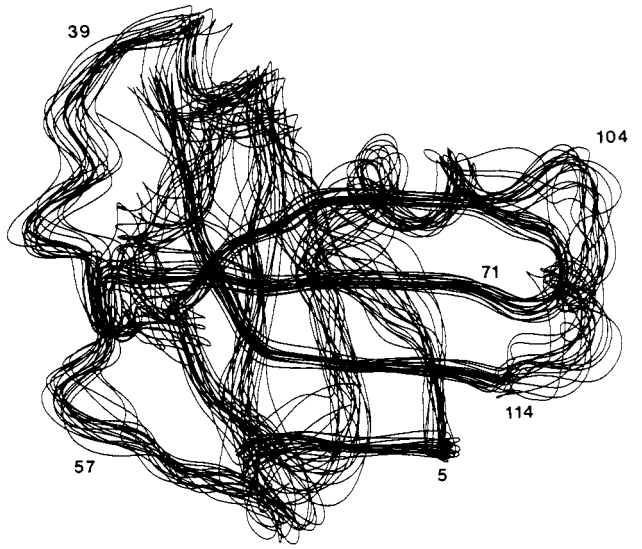

FIGURE 2: Twenty aligned restrained energy-minimized structures of barin represented by smoothed backbones of the atoms C^{α}, N, and C.

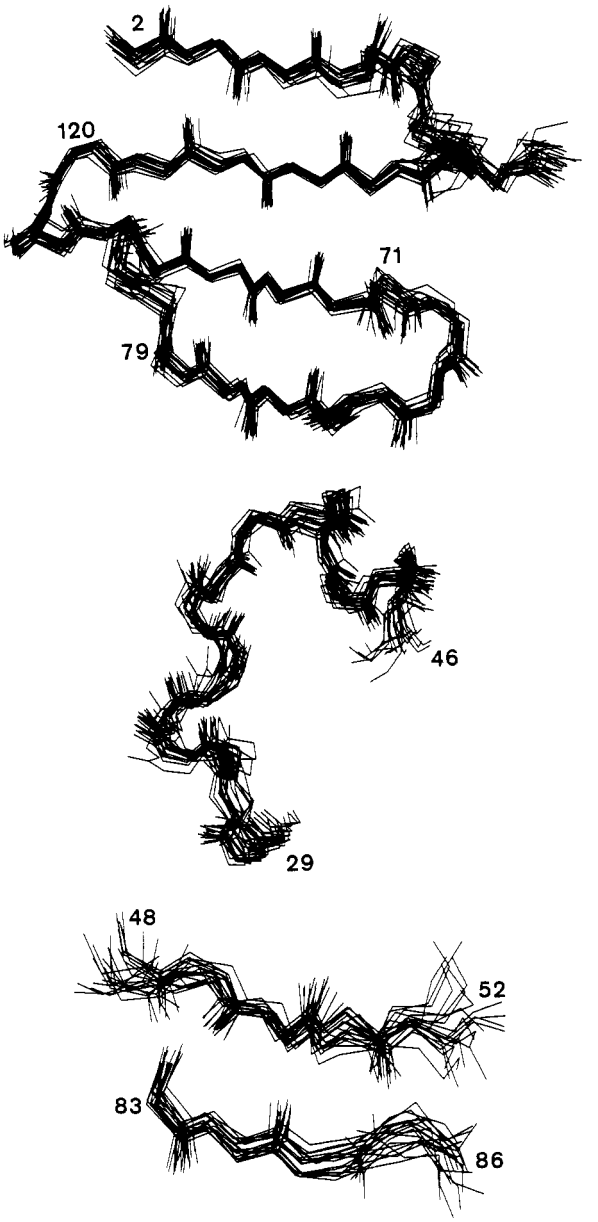

FIGURE 3: (a, top) Twenty aligned structures of the four-stranded antiparallel β-sheet including the loop of residues 72–75. Backbone atoms N, C^{α}, C, and O are shown. (b, middle) Backbone atoms of residues 29–46 consists of the two small α-helices α-I and α-II that bends at Pro39. (c, bottom) Backbone atoms of residues 48–52 (β-IIIb) and 83–86 (β-IVb).

The superposition of the backbone atoms of the 20 structures is shown in Figure 2. From the global alignment it is possible to follow the fold of the protein; however, alignment of selected regions shows that some of these are quite well-defined. Two regions of the protein (Figure 3a,b), the four-stranded antiparallel β-sheet, β-a, and the stretch from residue 29 to 46 including α-I and α-II, are individually well-defined with atomic RMS difference values of 0.37 and 0.59 Å, respectively, for backbone atoms. The α-helices α-III and α-IV, which both interact with the four-stranded antiparallel β-sheet, β-a, are, however, less well-defined with an atomic RMS difference value of 0.95 Å for the backbone atoms. The total atomic RMS difference for these well-defined regions of secondary structure β-a and α-1–α-4 is 0.89 Å for the backbone atoms. The parallel β-sheet (Figure 3c), β-IIIb and β-IVb, is also relatively well-defined with atomic RMS difference of 0.61 Å, whereas the alignment with strand β-IIb and β-Ib is poor due to the disorder of the residues from 89 to 95. The atomic RMS difference by residue after alignment of the entire backbone to the geometric average structure (Figure 4a) is not necessarily a good indication of the accuracy of the structure determination of different regions in a protein. Investigations of the backbone dihedral angles ϕ and ψ (Figure 4b,c) give a better picture of local geometric variations. ψ angles were better determined in regions with β-sheet structures due to the presence of strong sequential H^{α}–H^{N} (i,i+1) NOEs. ϕ angles of residue Val98 and Asp122 in the 20 structures are all close to the restraint limits, indicating stress in the structures at these two positions. In several of the glycine residues the dihedral angles were not well determined as a consequence of the lack of stereospecific assignment of the H^{α} protons.

Binding Site Studies. Perturbation of the specific protein resonance chemical shift values was only observed for the addition of the tetramer *N*-acetylglucosamine. For this experiment the perturbations of reasonably well resolved resonance lines of the protein were observed to influence only the chemical shift values, but no measurable change in line width was seen. The degree of change in chemical shift

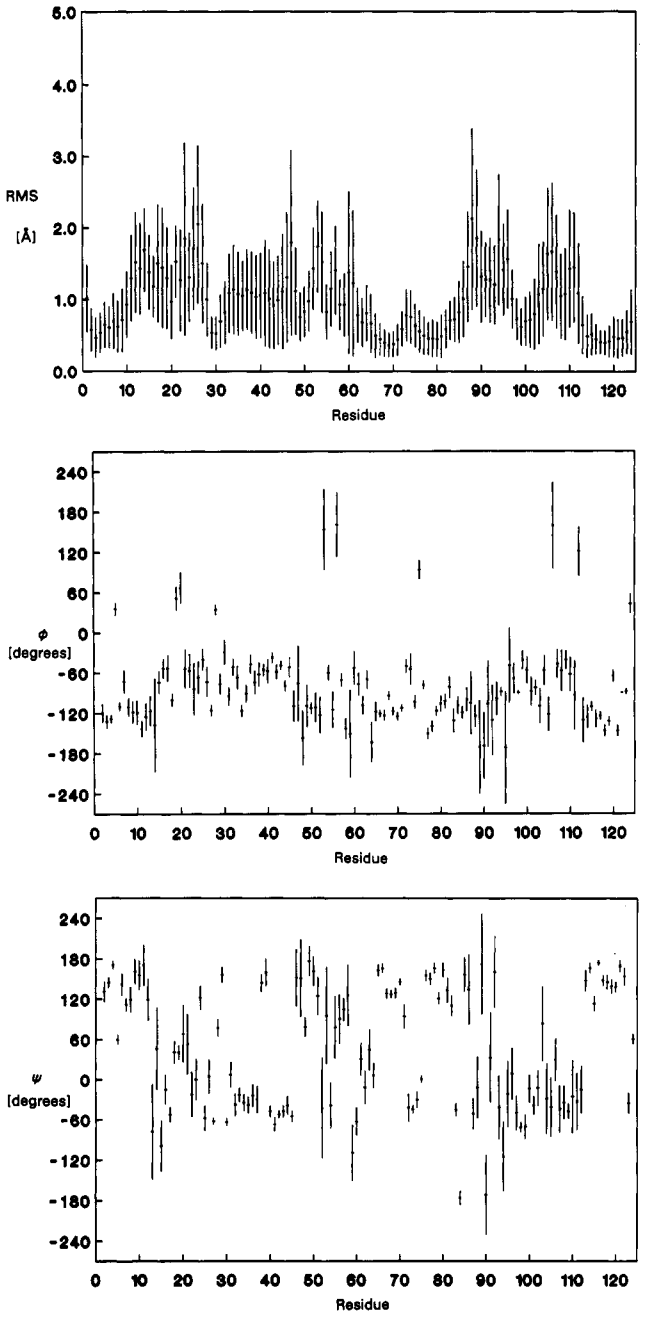

FIGURE 4: (a, top) Atomic RMS differences (ångström) by residue between each of the 20 structures and the mean structure. One standard deviation is shown as an error bar on each side of the mean value. (b, middle; c, bottom) Mean ϕ and ψ angles in degrees for the 20 structures of barwin, respectively, as a function of residue with error bars of two angular RMS values. Apart from Asp5, Asn19, Trp20, Ser28, and Arg124, which are known to have positive ϕ angles, Gly53, Gly56, Gly75, Gly106, and Gly112 have positive ϕ angles.

perturbation was influenced by the amount of added ligand. From these properties we concluded that the ligand was in fast exchange with the complex (Sudmeier et al., 1980). Since the protein concentration [P_{tot}] was constant and the total amount of added ligand [L_{tot}] was known, we could use the following equation to calculate the dissociation constant K_d assuming that only a single binding site existed:

$$[L_{tot}] = [P_{tot}]\frac{\delta_{obs} - \delta_P}{\delta_{PL} - \delta_P} + K_d\frac{\delta_{obs} - \delta_P}{\delta_{PL} - \delta_{obs}}$$

δ_{obs} is the observed chemical shift, δ_P is the chemical shift of the unperturbed resonance line of the protein, and δ_{PL} is the chemical shift of the same resonance line in the complex. K_d,

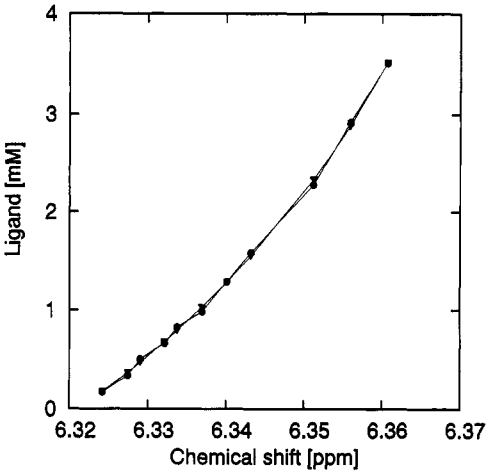

FIGURE 5: Concentration of the added tetramer *N*-acetylglucosamine vs chemical shift values for H^{α} of Tyr10. Solid circles are measured values, and hollow triangles are fitted data assuming a single binding site and fast exchange of ligand.

δ_P, and δ_{PL} were parameters determined using a least-squares fit with [L_{tot}] and δ_{obs} as variables. The measured as well as the fitted chemical shift values for the resonance line of H^{α} in Tyr10 are shown in Figure 5. K_d was determined to 5.2 ± 0.7 mM, which is characterized as a weak binding. Similar pertubations were observed for resonance lines of Tyr12, Leu93, Trp95, Tyr109, and Leu114 that were well resolved in the one-dimensional ^1H NMR spectra.

DISCUSSION

The structure of barwin (Figure 6) has an architecture of secondary structure composed of two β-sheet structures β-a and β-b, where the β-strands in the two sheets run essentially perpendicular to each other in two opposite shells of β structure. Inspection of the structures reveal that the C-terminal part of the (β-1a) including residues 7–9 and the N terminus of (β-IVa) including residues 113–115 are bent with respect to the remaining part of the other four antiparallel strands of βa. In the structure of barwin these two strands become an appendix to the β-b-sheet. The two sheets form the core of the protein and comprise 49/125 of the residues. Between the sheets are two hydrophobic regions: one centered around the aromatic residues Tyr118 and Phe120 and including the aliphatic residues Val5, Ala55, Leu67, Leu69, and Val116; the other hydrophobic region is centered around Trp20 and Trp48 and includes Ala16, Leu22, Thr49, Val83, and Leu93. The two β-sheets are flanked by two loop regions, L-I (Arg14–Ala29) and L-II (Gly53–Gly64), and by two strands of short α-helices, α-I and α-II between residues 31 and 46 and α-III and α-IV between residues 97 and 113. The two regions of α-helices are at opposite ends of the β-sheets. Both regions contain a bent helix consisting of two relatively short helices with axes almost perpendicular to each other. In the first stretch the helix is interrupted at Pro39 by a 3_{10} turn, and in the other stretch of helix the C-terminal part is a 3_{10}-helix including the two glycine residues (Gly106 and Gly108). Both helix regions contribute with side chains to the two hydrophobic cores of the structure, Leu40 and Trp42 to the first and Val98, Phe99, and Ile102 to the other region.

The three disulfide bridges in the structures connect in all cases regions of regular secondary structure. Cys31–Cys63 connects the C-terminal part of the first helix region (31–39) with the C-terminal part of strand III in the β-a-sheet. This strand is further linked to the C-terminal part of strand IV

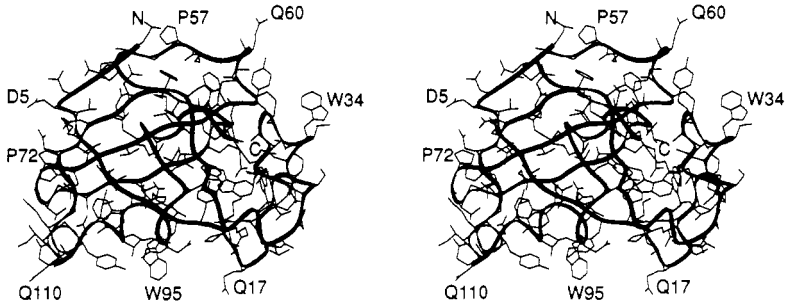

FIGURE 6: Stereoview of the solution structure of barwin. The peptide backbone is shown as a ribbon.

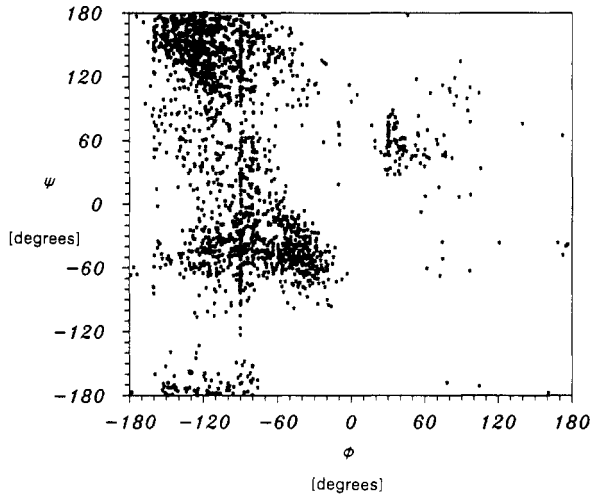

FIGURE 7: Distribution of ϕ, ψ angles (proline and glycine residues excluded) in all 20 structures in the Ramachandran plot (Ramachandran, 1968). Most ϕ, ψ points are in the sterically allowed regions; five residues have positive ϕ angles (Asp5, Asn19, Trp20, Ser28, and Arg124).

in the same sheet by the disulfide bond of residues 66 and 123. Strands III and IV of the β-sheet β-b are connected by the cystine bridge of residues 52 and 86. With this constellation of the three disulfide bridges the loop region L-II (Gly53–Gly64) is anchored to two regions of secondary structure.

One region located on the surface of the structure attracts some interest because it involves several tyrosyl, histidyl, tryptophanyl, and aspartate residues. Most of these groups are from strands I and II in sheet β-b and from the N-terminal part of the α-helix region of residues (97–113). In the NMR spectrum of this region a number of features were observed. First, the $H^{\delta 2}$ of His11 was observed at 4.97 ppm, almost 2 ppm lower than the typical chemical shift value. This low shift is explained by the position of the $H^{\delta 2}$ found in the structures right above the center of the phenol ring of Tyr13. Second, many of the NMR signals of strand β-IVb residues were broad and difficult to identify, as discussed in the previous paper (Ludvigsen & Poulsen, 1992). One of the residues is Trp95, which for these reasons was not sequence specifically assigned. Most of the resonances even in the indole ring are broad and reminiscent of the relatively broad ^{1}H NMR resonances of Trp62 in hen egg white lysozyme (Cassels et al., 1978), a residue in the binding site and close to the catalytic site in this enzyme.

The Ramachandron plot (Figure 7) of all 20 structures of barwin (glycine and proline excluded) shows that the majority of ϕ,ψ angles are in the geometrically allowed regions; however, the residues Asp5, Asn19, Trp20, Ser28, and Arg124 have all positive ϕ angles, in agreement with the observations of $^3J_{H^NH^\alpha}$ coupling constants and the intense intra H^N–H^α NOE effect.

Atomic RMS differences by residue of a global alignment of the N, C, and C^α in the peptide backbone to the average structure show as well that the β-sheet β-a is the best determined part of the structure with values in the range 0.30–0.50 Å (Figure 4a). The NMR data for this region are thus sufficient to define these regions. The remaining part of the peptide backbone is less well-defined with RMS differences in the range 0.8–1.5 Å for individual residues. A comparison of the individual ϕ and ψ angles in the structure (Figure 4b,c) shows that the angular RMS values are small for most residues, even for those where the atomic RMS differences were relatively high. This implies that the dihedral angle restrains and intraresidue and sequential NOEs in most cases were sufficient to define the backbone conformation of the individual residues but the number of long-range NOEs was not sufficient to determine the global geometry as accurately.

In the ^{1}H NMR spectra of barwin there were clear indications that certain regions of the protein accommodate at least two distinguishable conformations in slow exchange and that other parts had very broad resonances indicating structural mobility. Indeed, from the secondary structure evaluation all of the hydrogen bonds that would define sheet β-b were not observed. Only when the structure calculation had shown the arrangements of the four strands did it become clear that the four-stranded sheet could exist. The slowly exchanging amides observed for residues His11, Tyr13, Trp48, Ala50, Cys52, and Leu93 in the strands clearly suggested the presence of hydrogen bonds that were not inferred by cross-strand NOEs that are typically encountered in β-sheets. We propose that the lack of these NOEs is due to dynamics in this region of the structure possibly imposed by cis/trans isomerism of the Pro15 and Pro54. However, it should be emphasized that no direct evidence for a cis/trans isomerism of either Pro15 or Pro54 has been observed in the NMR spectra unambiguously.

Inspection of the structures revealed that a number of amide hydrogens that were identified as slowly exchanging with solvent were not engaged in hydrogen bonds in the calculated structures. Studies of the correlation between hydrogen bonding and slow amide hydrogen exchange with solvent (Pedersen et al., 1991) have shown that slow amide exchange is essentially always associated with hydrogen bond formation in a protein of this size. This implies that the amides with slow exchange and no hydrogen bonds formed in the calculated structures may indeed form hydrogen bonds. However, these apparently could not be defined by the NMR restraints applied in the structure calculation. Given this information, it was of interest to examine whether the combined information of the main globular fold and the specific sites of slow amide hydrogen exchange could reveal hydrogen bonds in the structures that could not be inferred from the NMR study directly. Therefore, the sites were examined for potential

hydrogen bond acceptors in the proximity of the amide hydrogen. These hydrogen bonds are in addition to the ones already listed in Table I (from H^N to CO, residues 11–90, 13–92, 92–11, 94–13, 52–91, 93–50, and 50–93). To examine whether these hydrogen bonds can exist in the framework of structures defined by the NMR restraints, a new structure calculation was performed where these hydrogen bonds were added as restraints to the existing list of restraints and a new structure calculation was performed using the substructures obtained previously from DISGEO. These new structures came out with NOE energies on the average of 80 kcal/mol, slightly higher than for the original structure, however, still low. This assembly of structures has an atomic RMS difference of 1.12 Å for the backbone atoms.

Given the amino acid sequence similarity, the three closely related wound-induced proteins from potato and rubber are likely to have a globular fold similar to that of barwin. Thus, it is of interest to compare the structures determined here with the amino acid sequences for these other three proteins, in particular, to examine the sites in the structures where the sequence identity was conserved in all proteins and similarly the sites where mutations were observed. The comparison shows that the sites in the sequence where mutations were most frequent are on the surface of the protein:in the loop region between Pro15 and Pro25, where only the Trp, Asp, Leu sequence is conserved; in the loop region between Pro54 and Gly64, where still the glycine residues are all preserved; and in the helix region between Trp95 and Gly111. None of the cysteine residues are mutated. In the hydrophobic regions between the two β-sheets only conservative mutations have occurred. It can be noted that the glycine residues throughout these sequences are strongly conserved. The mutation frequency of glycine residues is commonly high; however, in this group of sequences they are strongly preserved. This suggests that the glycine residues in these proteins structures are important to maintain the folding of the peptide chain, probably by turn formation.

The binding of the β(1-4) tetramer of *N*-acetylglucosamine in barwin was found to be relatively weak. However, the finding that neither *N*-acetylglucosamine monomer nor any of the three disaccharides maltose, chitobiose, and lactose gave any chemical shift perturbation of protein resonances suggests that barwin has no affinity for these. The affinity for the *N*-acetylglucosamine oligomer, therefore, is not associated with any stronger binding of any of these disaccharides. The binding site for the oligosaccharide has been localized by the chemical shift perturbation. These are observed for resonances of residues in one particular region of the protein. This region is formed by the β-b-sheet by the residues on the surface of this sheet. It consists of three tyrosine residues, Tyr10, Tyr12, and Tyr109, two histidine residues, His11 and His113, two aspartate residues, Asp92 and Asp94, and Trp95 present in a confined region on the protein surface forming a potential binding site.

The topology of the structure of barwin is not common, and so far we are not aware of any other protein with a similar topology. However, at present we are not in position to claim that barwin represents a new topology of protein structures.

In conclusion, the determination of the three-dimensional structure of barwin has revealed the overall global fold of the peptide backbone, built over a hydrophobic core shelled by the secondary structure elements. Furthermore, as a result of the work it has been possible to identify one site on the surface of the protein that could be a potential binding site. Finally, it could be shown that the three-dimensional structure determined here in terms of the global fold and secondary structure is most likely similar to the protein domains of the related proteins encoded by wound-induced genes.

ACKNOWLEDGMENT

We thank Pia Mikkelsen for isolation and purification of barwin and Mogens Kjær for help and advice with the structure calculations. We also thank Jens Chr. Madsen and Kim V. Andersen for a critical reading of the manuscript. F.M.P. is a member of the Danish Protein Engineering Centre (PERC).

REFERENCES

Broekaert, W., Lee, H., Kush, A., Chua, N.-H., & Raikhel, N. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 7633–7637.

Brooks, B. R., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

Brünger, A. T. (1988) X-PLOR software, Yale University and Harvard University.

Cassels, R., Dobson, C. M., Poulsen, F. M., & Williams, R. J. P. (1978) *Eur. J. Biochem. 92*, 81–97.

Clore, G. M., Gronenborn, A. M., Brünger, A. T., & Karplus, M. (1985) *J. Mol. Biol. 185*, 435–455.

Clore, G. M., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986) *J. Mol. Biol. 191*, 523–551.

Crippen, G. M., & Havel, T. F. (1978) *Acta Crystallogr. A34*, 282–284.

Gelin, B. (1976) Ph.D. Thesis, Harvard University.

Günthert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517–530.

Havel, T. F. (1986) *QCPE 507*.

Havel, T. F., & Wüthrich, K. (1984) *Bull. Math. Biol. 46*, 673–698.

Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665–720.

Havel, T. F., Kuntz, I. D., & Wüthrich, K. (1985) *Bull. Math. Biol. 45*, 665–720.

Kline, A. D., Braun, W., & Wüthrich, K. (1988) *J. Mol. Biol. 204*, 675–724.

Ludvigsen, S., & Poulsen, F. M. (1992) *Biochemistry* (second of three papers in this issue).

Ludvigsen, S., Andersen, K. V., & Poulsen, F. M. (1991) *J. Mol. Biol. 217*, 731–736.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988) *FEBS Lett. 2*, 317–324.

Pedersen, T. G., Sigurskjold, B. W., Andersen, K. V., Kjær, M., Poulsen, F. M., Dobson, C. M., & Redfield, C. (1991) *J. Mol. Biol. 218*, 413–426.

Powell, M. J. D. (1977) *Math. Prog. 12*, 241–254.

Ramachandran, G. M., & Sasisekharan, V. (1968) *Adv. Protein Chem. 23*, 283–437.

Sippl, M. J., & Scheraga, H. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2283–2287.

Stanford, A., Bevan, M., & Northcote, D. (1989) *MGG, Mol. Gen. Genet. 215*, 200–208.

Sudmeier, J. L., Evelhoch, J. L., & Jonsson, N. B.-H. (1980) *J. Magn. Reson. 40*, 377–390.

Svensson, B., Svendsen, I., Højrup, P., Roepstorff, P., Ludvigsen, S., & Poulsen, F. M. (1992) *Biochemistry* (first of three papers in this issue).

Venkatachalam, C. M. (1968) *Biopolymers 6*, 1425–1436.

Verlet, L. (1968) *Phys. Rev. 159*, 98–105.

Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295–315.

Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.