^{1}H NMR Assignments and Secondary Structure of Human β_2-Microglobulin in Solution

Mark Okon‡

Institute of Biological Physics, 142292 Pushchino, Russia

Paul Bray

Institute of General and Physical Chemistry, Belgrade University, 11000 Belgrade, Yugoslavia

Dušan Vučelić*

Faculty of Sciences and Mathematics, School of Physical Chemistry, Belgrade University, Studentski trg 16, 11000 Belgrade, P.O. Box 550, Yugoslavia

Received February 7, 1992; Revised Manuscript Received June 1, 1992

ABSTRACT: Sequence-specific resonance assignments of human β_2-microglobulin (M_r 12 000) and its secondary structure are determined by 2D NMR techniques. The protein is found to contain two antiparallel β-sheets each of four β-strands with the β-sheets being connected by a single disulfide linkage. No evidence for any regular helical structure is found. Amide proton-solvent-exchange rate constants and $^3J_{HN\alpha}$ coupling constants are evaluated.

β_2-Microglobulin (β_2m),[1] a small protein (12 000 MW), has a crucial role in immunology. This protein occurs as a monomer at relatively low concentrations (around 1.8 mg/L) in biological fluids such as serum, cerebral spinal fluid, and colostrum (Berggard & Bearn, 1968). β_2m is the small invariant subunit of the highly polymorphic class I histocompatibility (human HLA) antigens which are located on the external surface membrane and within all nucleated mammalian cells (Grey et al., 1973; Peterson et al., 1974; Ploegh et al., 1981a). Internal HLA antigens could be either biosynthetic precursors enroute to the surface membrane or, alternatively, internalized surface antigens (Ploegh et al., 1981b; Tse & Pernis, 1984).

The role of the kidneys in catabolism of small blood proteins by glomerular filtration and tubular degradation is very well known. Serum levels of β_2m are an indicator of renal glomerular function since small proteins are removed from blood by glomerular filtration. Thus, serum β_2m concentration is inversely correlated with glomerular filtration rate (Wibell et al., 1973; Trollfors & Norrby, 1981). Urinary β_2m levels in healthy subjects are very low in the range of 0.06–0.17 mg per 24-h urine volume (Berggard & Bearn, 1968; Peterson et al., 1969). High levels of urinary β_2m are an indication of renal tubular malfunction with causes such as renal transplant rejection, cadmium toxicity, aminoglycoside nephrotoxicity, and endemic Balkan nephropathy (Berggard & Bearn, 1968; Peterson et al., 1969; Hall & Vasiljević, 1973; Karlsson & Lenkei, 1977).

Amyloidosis is a group of diseases characterized by fibrillar protein deposits in various tissues. Several proteins have been found to form amyloid fibrils [Bence-Jones proteins in multiple myeloma (Tveteraas et al., 1985), variant prealbumin in polyneuropathy of Jewish origin (Pras et al., 1983), and a Swedish form of this latter disease (Skinner & Cohen, 1981)]. Recent evidence suggests a direct pathological role for β_2m since amyloid deposits in chronic hemodialysis contain intact β_2m as a major component (Gejyo et al., 1985; Gorevic et al., 1985). Increased plasma β_2m concentrations are chronic in this condition (Vincent et al., 1978; Wibell et al., 1973) since hemodialysis often does not efficiently remove β_2m from plasma.

β_2m was originally isolated from urine of patients with tubular proteinuria and characterized by Berggard and Bearn (1968). The amino acid sequence of β_2m established a basic structure related to immunoglobulin constant domains (Smithies & Poulik, 1972; Cunningham et al., 1973). Indeed, β_2m is as closely related to the CL, CH1, CH2, and CH3 constant domains of immunoglobulin G as these constant domains are related to each other. Thus, a β_2m is a simple model with which to study an immunoglobulin structure since it occurs in biological fluids as a single-domain monomer. Immunoglobulin domains have two antiparallel β-sheets. Thus, β_2m might be amenable to conformational rearrangements since antiparallel β-sheets can accomodate a variety of conformations by variations in twisting, coiling, and extension of β-strands (Chothia, 1973; Salemme & Weatherford, 1981; Salemme, 1983; Chothia & Janin, 1981).

The features of main-chain folding of β_2m have been determined in crystals by X-ray crystallography of bovine β_2m (Becker & Reeke, 1985) and in more detail for human class I histocompatibility antigens, HLA-A2 and HLA-Aw68.1 (Bjorkman et al., 1987a,b).

In this paper, we report proton NMR assignments of human β_2m, secondary structure specification, β-strand alignments and some side-chain interactions.

* To whom correspondence should be addressed.

‡ Present address: Institute of General and Physical Chemistry, Belgrade University, 11000 Belgrade, Yugoslavia.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; 2D, two dimensional; COSY, 2D correlated spectroscopy; RECOSY, 2D relayed-coherence-transfer spectroscopy; NOESY, 2D nuclear Overhauser enhancement spectroscopy; CONOE, combined COSY-NOESY experiment; TOCSY, 2D total-correlated spectroscopy; DSS, sodium 4,4-dimethyl-4-silapentane-1-sulfonate; β_2m, β_2-microglobulin; pH*, pH meter reading without correction for the deuterium isotope effect.

0006-2960/92/0431-8906$03.00/0

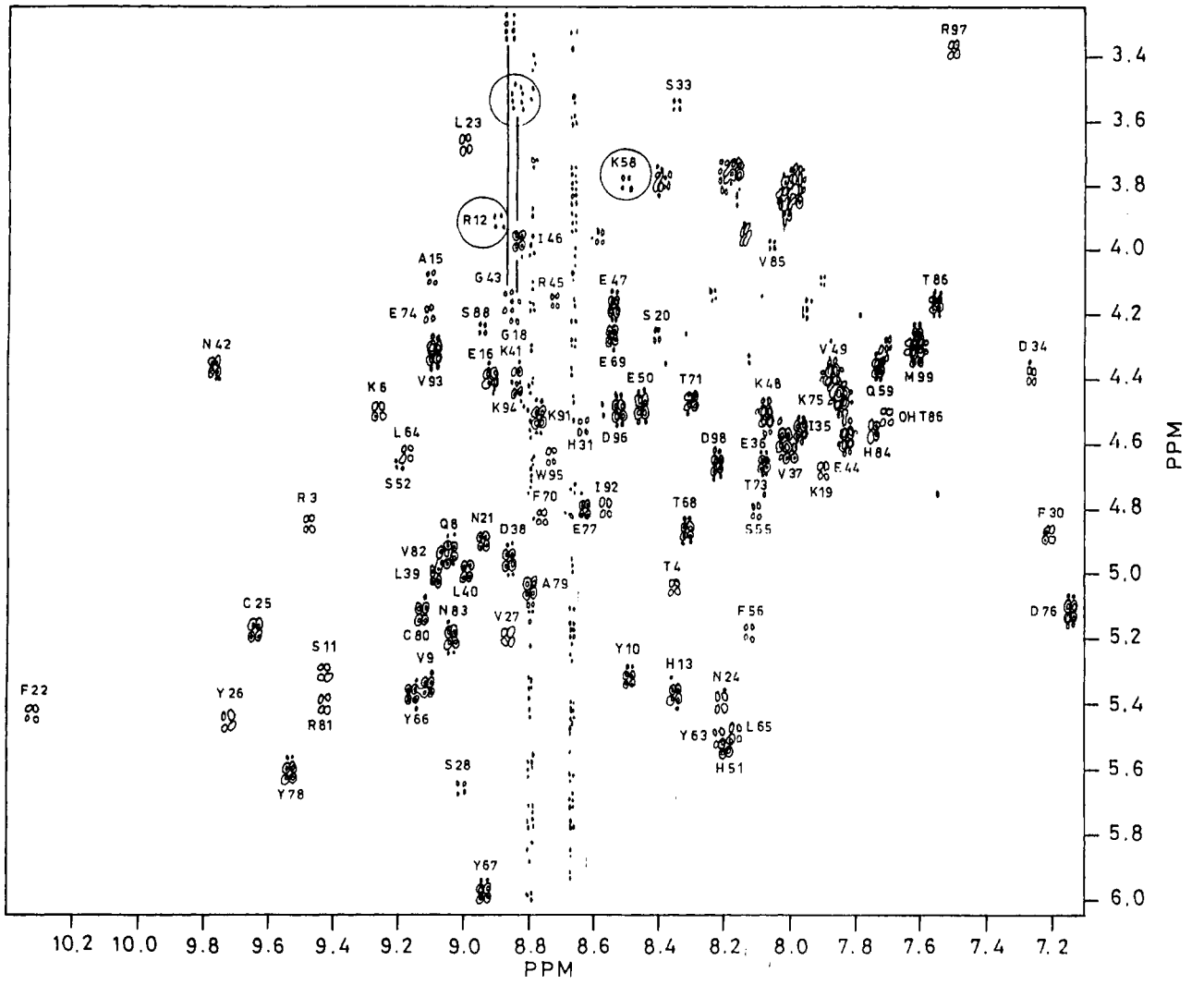

FIGURE 1: Fingerprint region of the phase-sensitive 600-MHz COSY spectrum of human β_2m in H_2O. The concentration of β_2m is 6 mg/mL, at pH 6.0, 30 °C. NH–αCH cross-peaks observed at this pH are seen. Cross-peaks in circles are plotted at a lower level.

MATERIALS AND METHODS

Sample Preparations. β_2m was purified from urine of a single patient in rejection crisis after renal transplantation or from Balkan nephropathy patients using a modification of the procedure of Berggard and Bearn (1968). Urine was kept at 0–10 °C during each 24-h collection period being collected in 1-L bottles containing 50 mL of 1.0 M Tris-HCl buffer, pH 7.8, with 4 mM sodium cocodylate to neutralize pH and inhibit bacterial growth. Purification of β_2m was achieved on columns of Sephadex G-150-40 (two times), DEAE-Sephacel, and Sephadex G-75-40. Purity of β_2m was established by single bands after electrophoresis in slab gels in presence of SDS at acid pH, in absence of SDS at pH 8.8, and by isoelectric focusing.

The samples for NMR in 95% H_2O/5% D_2O or 99.8% D_2O consisted of 0.3–1 mM β_2m, 150 mM NaCl, and 50 mM phosphate buffer adjusted to pH 5.4–8.0. Temperature was varied from 22 to 47 °C.

NMR Spectroscopy. NMR spectra were acquired with Bruker WM 400 or AM 600 spectrometers. ^{1}H chemical shifts are reported relative to internal DSS at 0 ppm.

Phase-sensitive 2D COSY spectra (Marion & Wüthrich, 1983) of β_2m in H_2O were acquired with a D_1–90°$_\phi$–t_1–90°$_\psi$–(+Acq)–D_1–180°–90°$_\phi$–t_1–90°$_\psi$–(–Acq) pulse sequence, where D_1 is the relaxation delay time and Acq is the acquisition period. This implementation of a 180° pulse in the pulse sequence reduces the residual water signal by more than a factor of 5.

Phase-sensitive 2D NOESY spectra (Kumar et al., 1980; Bodenhausen et al., 1984) were collected in an alternated mode with COSY spectra to prevent any influence of time. 2D CONOE experiments (Gurevich et al., 1984) were in D_2O as solvent. In some NOESY experiments, a 1–$\bar{1}$ pulse (Hore, 1983) was used for the third pulse in the standard NOESY pulse sequence. A mixing time (τ_m) of 200 ms was used in all NOESY experiments with no random variations. Phase-sensitive 2D RECOSY (RELAY) spectra in D_2O (Wagner 1983; Bax & Drobny, 1985) were recorded with a 30-ms mixing time.

Clean TOCSY experiments (Griesinger et al., 1988) were recorded with a mixing time τ_m = 60 ms, a trim pulse of 2 ms, and $r = \Delta/\tau_{90°} = 2.5$ ($\tau_{90°}$ is duration of a 90° pulse, and Δ is the delay time before and after 180° pulse in the MLEV-17 sequence). In all experiments in H_2O, the water signal was saturated during relaxation time D_1 = 1 s except in NOESY experiments with a 1–$\bar{1}$ pulse. Data were collected using 512 values of t_1; 2048 points were collected in t_2. Typically, 256 scans were collected for each t_1 increment. Before 2D Fourier transformation, the time domain matrix was expanded to 2048 points in t_1 and 4096 points in t_2 by zero-filling. COSY and RECOSY spectra were processed

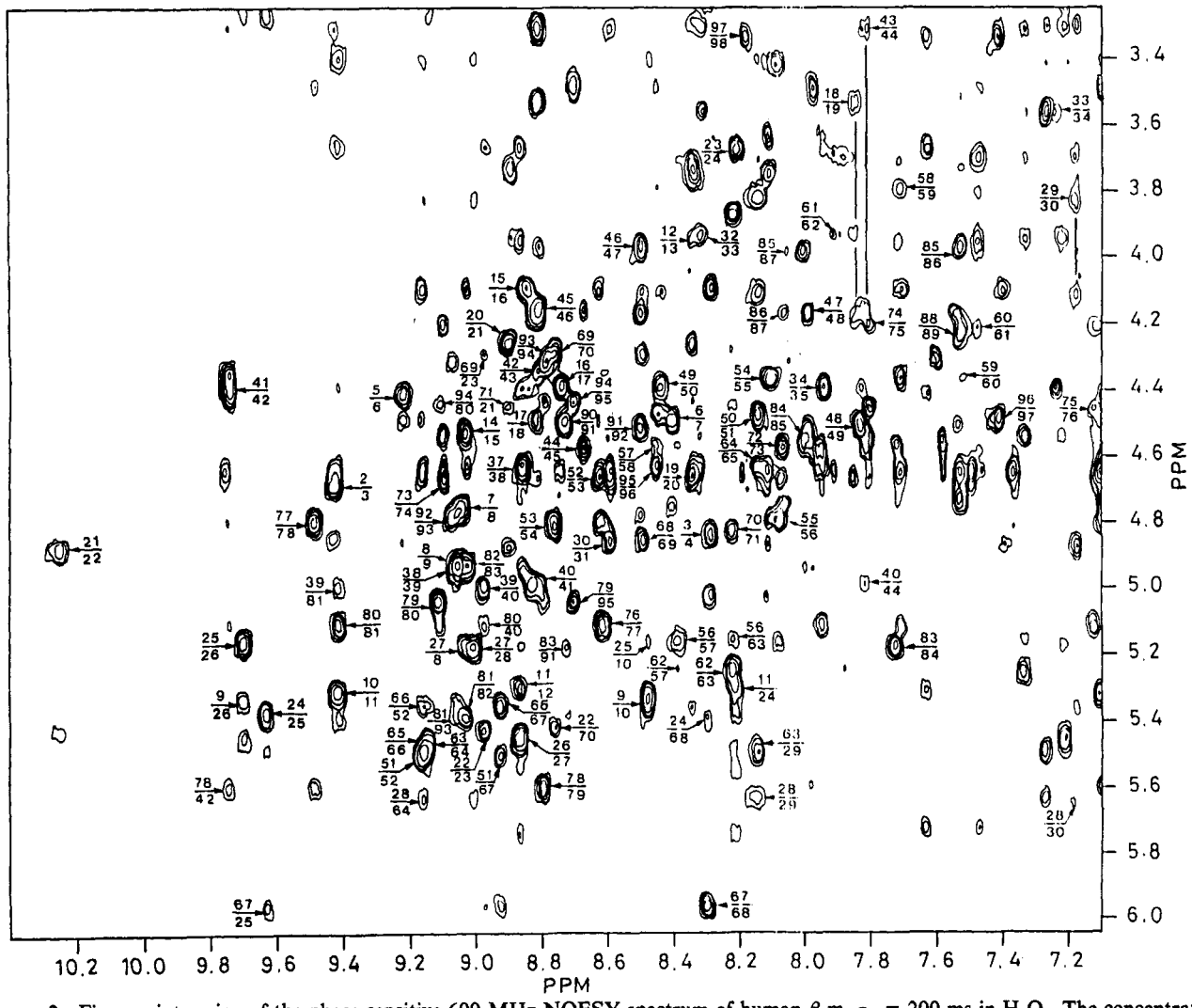

FIGURE 2: Fingerprint region of the phase-sensitive 600-MHz NOESY spectrum of human β_2m, τ_m = 200 ms in H_2O. The concentration of β_2m is 8 mg/mL, at pH 5.9, 37 °C. Sequential $d_{\alpha N}(i,i+1)$ as well as middle $d_{\alpha N}(i,i+2)$ and long-range $d_{\alpha N}(i,j)$ connectivities identified in the assignment process are labeled.

with sine bell windows without phase shifts. NOESY and TOCSY spectra were processed with phased-shifted sine bell windows in D_2O experiments and with a trapezoidal window in t_2 dimension and a phase-shifted sine bell window in t_1 dimension in H_2O experiments.

In evaluating $^3J_{HN\alpha}$ values, 512 increments with 4K data points and 7246-Hz spectral width were used in the COSY experiment. Final digital resolution in the F_2 dimension was 0.2 Hz. Coupling constants were measured as a described by Marion and Wüthrich (1983). Real values of $^3J_{HN\alpha}$ are somewhat smaller than measured values, especially for ones <6 Hz, because accuracy is adversely affected by the relatively large line widths (>10 Hz) of amide proton resonances (Neuhaus et al., 1985).

NH Exchange Rates. Slowly exchanging amide protons were observed as cross-peaks in four COSY spectra, each taking 4 h at pH 6.0 and 27 °C immediately after dissolving lyophilized β_2m powder in D_2O. Spectral widths were 6.25 and 2.62 KHz in F_2 and F_1, respectively. A total of 200 t_1 increments of 1K data points were recorded with 64 scans for each t_1 increment.

Crystal Structure. Atomic coordinates and secondary structure specifications for histocompatibility antigens HLA-A2.1 and HLA-Aw68.1 (Bjorkman et al., 1987a,b: Garrett et al., 1989) were obtained from the Brookhaven National Laboratory Protein Data Bank, Upton, New York (Bernstein et al., 1977).

RESULTS AND DISCUSSION

Sequence-Specific Assignments

Assignment of the proton spectrum of human β_2m was done using a standard two-step approach: establishment of "spin systems" and "sequential assignments" (Billeter et al., 1982; Wüthrich, 1986).

From 2D COSY, NOESY, and RELAY spectra of β_2m in D_2O, 36 spin systems were obtained in a straightforward manner including all spins from Thr, Val, Ala, four Ile up to γCH_3, one Leu (23), and all ring protons (His, Tyr, Phe, and Trp).

Sequential assignments were made from COSY and NOESY spectra of β_2m in H_2O recorded at different pH and temperatures. Fingerprint regions of COSY (at pH 6.0 and 30 °C) and NOESY (at pH 5.9 and 37 °C) spectra are shown in Figures 1 and 2, respectively. Sequence-specific assignments for extended peptide segments were made on the basis of strong $d_{\alpha N}(i,i+1)$ NOE connectivities. In this way, segments 8–12, 20–28, 37–41, 46–51, 62–71, 75–84, and 91–95 were obtained. Each segment consists of one to six spin systems assigned from previous experiments in D_2O. Typical examples of these

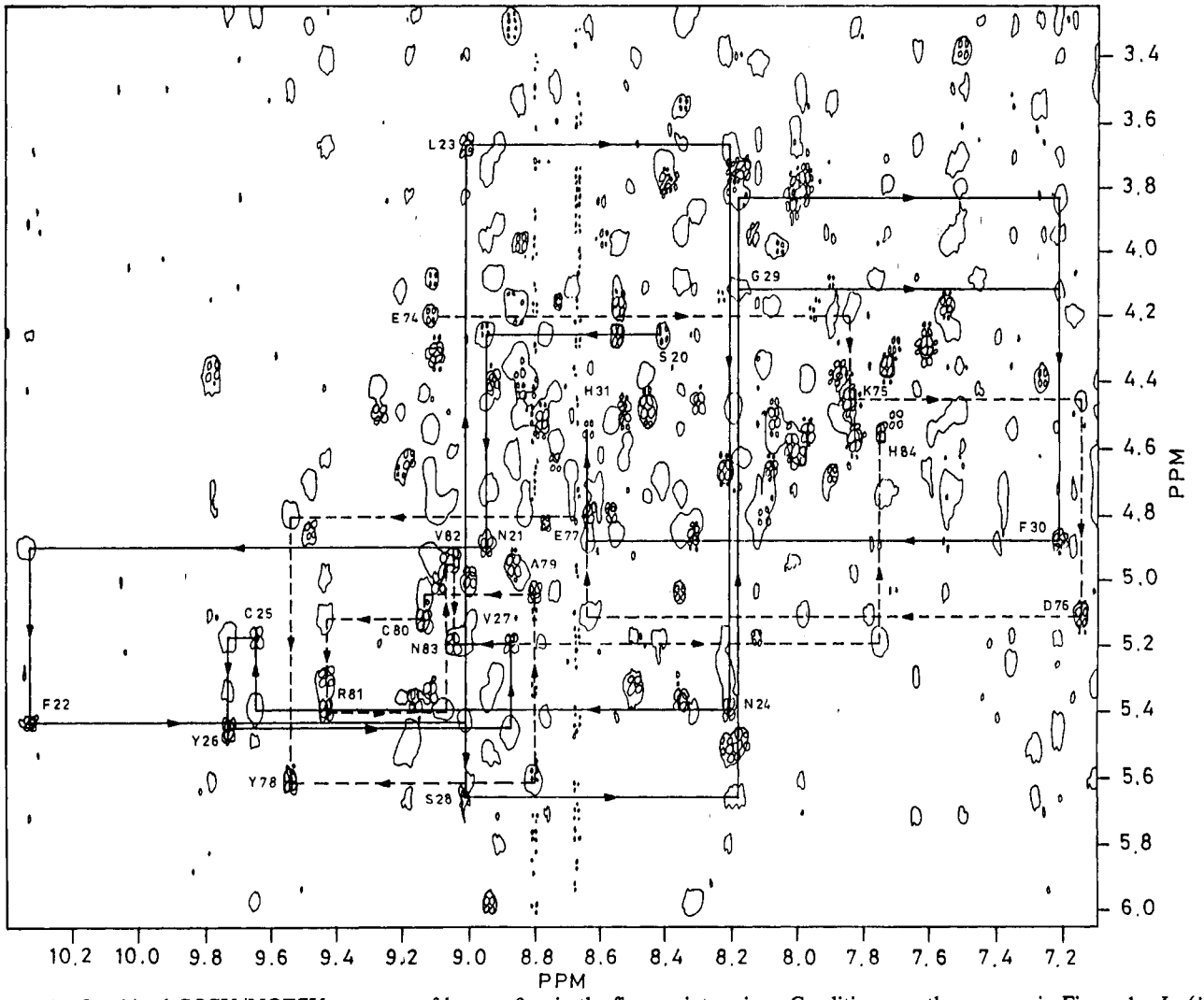

FIGURE 3: Combined COSY/NOESY spectrum of human β_2m in the fingerprint region. Conditions are the same as in Figure 1. $J_{N\alpha}(i,i)$ and sequential $d_{\alpha N}(i,i+1)$ connectivities for residues 20–31 and 74–84 are shown. Only one level is drawn.

assignments are given for segments 20–31 and 74–84 in Figure 3. The combined $J_{N\alpha}(i,i)$–$d_{\alpha N}(i,i+1)$ connectivity sequence from the segment 20–31 includes Leu and Val, whose spin systems were assigned from D_2O spectra, in respective sequence positions i and $i+4$. The only unique possibility which corresponds in the amino acid sequence is fragment –L23–N24–C25–Y26–V27–. Similarly, segment 74–84 includes spin systems Ala(i) and Val($i+3$) giving –A79–C80–R81–V82–.

To solve the problem of amino acid side-chain assignments, TOCSY experiments were used being supported by COSY, RELAY, and NOESY both in H_2O and D_2O solutions. TOCSY experiments uniquely provided NH-ϵCH_2 connectivities in some Lys (K19, K48, and K75) and αCH-δCH_3 connectivities in some Leu and Ile (I7, L39, L40, L65, L87, and I92). The high dispersion of the NH and αCH shifts (10.39–7.13 and 5.98–3.42 ppm, respectively) enabled separate observation of the entire spin systems of individual residues. It solved the problem of overlapping resonances in aliphatic region. Nevertheless, although TOCSY spectra in D_2O and H_2O with τ_m = 60 ms contain much information, in this paper they are used only for assignment of a few signals from Ser 57, and 61, all Lys (except K41), all γCH_2 of Gln and Glu, and connections from βCH_2 to γCH for all Leu except L23 and from βCH to γCH_2 for all Ile except I1. All other individual connectivities were confirmed by COSY experiments.

Aromatic protons were assigned by intraresidue NOE connectivities between βCH_2 and ring protons: C2,6H for Tyr and Phe, C2H of Trp or C4H of His. NOE's between βCH_2 and δNH_2 for Asn and γCH_2 and ϵNH_2 for Gln were also used for assignments.

Arg assignments were made up to ϵNH from COSY spectra, but ϵNH of R12 is still missing.

On the basis of this approach, signal identifications were established for segments 96–99, 42–45, 13–16, 18–19, 34–36, 73–74, 6–7, 29–31, 85–87, 52–56, 88–89, 3–4, 58–61, and 17. Pro (5, 14, 32, 72, 90), S57, S33, Q2, and I1 were the last to be assigned. The amide (NH) proton of Q2 was not found, and consequently $d_{\alpha N}$ between I1 and Q2 is missing. Ile 1 behaves as an amino terminus with sharp peaks and an αCH chemical shift dependent upon pH, ionic strength, and temperature.

At standard experimental conditions of 37 °C, 7×10^{-4}M β_2m, and pH 6.0, line widths of peptide NH's are 13–15 Hz and ~10 Hz for methyl groups. Consequently, a high signal/noise ratio in "fingerprint region" can be expected for residues with large $^3J_{HN\alpha}$ in COSY spectra and for short sequential $d_{\alpha N}$ distances in NOESY spectra, both being typical for residues in β-structures. Signals with small $^3J_{HN\alpha}$ below 6 Hz usually are very weak in COSY spectra. For example, in COSY spectra in Figure 1, resonances from A15, H31, S33, K58, E74, and V85 are very weak at 600 MHz and completely

Table I: ^{1}H Assignments and Chemical Shift Data for Human β_2m at pH 7.0, 37 °C[a]

residue	NH	αCH	βCH	γCH and others
Ile 1		3.78	1.78	γCH 1.35, 1.07; γCH_3 0.82; δCH_3 0.77
Gln 2		4.70	2.13, 2.08	γCH 2.54, 2.40; εNH 7.50, 6.74
Arg 3	9.41	4.83	2.23, 1.84	γCH 1.75, 1.68; δCH 3.20, 3.11; εNH (7.14)
Thr 4	8.28	5.03	4.09	γCH_3 1.32
Pro 5		4.42	1.55	γCH 1.68, 1.14; δCH 3.67, 3.15
Lys 6	9.23	4.51	1.77, 1.50	γCH 1.63, 1.42
Ile 7	8.41	4.78	1.64	γCH 1.45; γCH_3 0.77; δCH_3 0.73
Gln 8	9.04	4.93	2.35, 2.20	γCH 2.41, 2.41; εNH 7.46, 6.65
Val 9	9.07	5.35	2.03	γCH_3 0.97, 0.91
Tyr 10	8.48	5.32	3.41, 3.19	C2,6H 6.82; C3,5H 6.56
Ser 11	9.35	5.29	4.42, 3.66	
Arg 12	8.88	3.92	1.93, 1.74	γCH 1.43; δCH 3.05, 2.78
His 13	8.32	5.34	3.22, 2.89	C2H 7.84; C4H 7.07
Pro 14		4.53	2.43, 1.94	γCH 2.24, 2.08; δCH 4.08, 3.80
Ala 15	9.02	4.10	1.67	
Glu 16	8.87	4.40	1.85	γCH 2.15, 2.11
Asn 17	(8.79)	4.50	2.78, 2.68	δNH 7.53, 7.12
Gly 18	8.84	4.20, 3.53		
Lys 19	7.85	4.67	1.82, 1.74	γCH 1.35, 1.29; δCH 1.63, 1.58; ϵCH_2 2.93
Ser 20	8.36	4.25	3.76, 3.73	
Asn 21	8.91	4.88	2.73, 2.52	δNH 7.72, 7.46
Phe 22	10.39	5.43	2.69, 2.62	C2,6H 7.01; C3,5H 7.39; C4H 7.32
Leu 23	8.99	3.68	0.80, –0.82	γCH 0.64; δCH_3 0.01, –0.58
Asn 24	8.20	5.39	1.83, 1.44	δNH 5.74, 5.53
Cys 25	9.63	5.17	3.28, 2.54	
Tyr 26	9.69	5.44	3.22, 3.20	C2,6H 7.21; C3,5H 6.65
Val 27	8.88	5.18	1.92	γCH_3 0.91, 0.76
Ser 28	9.00	5.64	3.81, 3.41	
Gly 29	8.16	4.12, 3.81		
Phe 30	7.19	4.87	2.32, 1.98	C2,6H 7.00; C3,5H 7.26; C4H 6.95
His 31	8.60	4.54	3.31, 3.13	C2H 7.87; C4H 7.05
Pro 32		3.90	2.34	γCH 2.03, 1.70; δCH 3.75, 3.65
Ser 33	8.29	3.54	3.31, 1.86	
Asp 34	7.24	4.40	2.43, 2.34	
Ile 35	7.96	4.56	1.40	γCH 1.50, 0.67; γCH_3 0.58; δCH_3 –0.51
Glu 36	7.98	4.59	1.91, 1.73	γCH 2.08, 2.03
Val 37	7.98	4.63	0.45	γCH_3 0.48, 0.24
Asp 38	8.86	4.96	2.38, 2.17	
Leu 39	9.06	5.01	1.66, 1.23	γCH 1.61; δCH_3 0.76, 0.74
Leu 40	8.98	5.00	1.56, 0.82	γCH 1.19; δCH_3 0.57, 0.40
Lys 41	8.84	4.41	1.72, 1.42	γCH 0.80, 0.56; δCH 1.75, 1.61; εCH 2.83, 2.79
Asn 42	9.73	4.37	2.92, 2.88	δNH 7.98, 7.72
Gly 43	8.80	4.17, 3.32		
Glu 44	7.84	4.56	2.03, 1.93	γCH 2.27, 2.15
Arg 45	8.68	4.18	1.64, 1.64	γCH 1.52, 1.34; δCH 3.12, 3.05; εNH 7.38
Ile 46	8.81	3.98	1.53	γCH 1.69, 1.02; γCH_3 0.96; δCH_3 0.85
Glu 47	8.49	4.18	2.12, 2.02	γCH 2.34, 2.22
Lys 48	7.98	4.51	1.93, 1.78	γCH 1.43, 1.38; δCH_2 1.71; ϵCH_2 3.02
Val 49	7.83	4.40	2.11	γCH_3 1.08, 1.02
Glu 50	8.46	4.47	1.56, 0.75	γCH 2.06, 2.06
His 51	8.10	5.50	2.47, 2.04	C2H 8.44, C4H 6.99
Ser 52	9.13	4.67	4.48, 4.09	
Asp 53	(8.66)	4.81	2.76, 2.57	
Leu 54	(8.80)	4.37	1.84	γCH 1.70; δCH_3 1.05, 0.86
Ser 55	8.10	4.78	3.43, 2.79	
Phe 56	8.08	5.14	2.62, 2.62	C2,6H 6.41; C3,5H 6.98; C4H 7.04
Ser 57	8.35	4.56	3.93, 3.68	
Lys 58	(8.49)	3.80	1.73	γCH_2 1.37; δCH_2 1.61; ϵCH_2 2.94
Asp 59	7.71	4.36	2.30	
Trp 60	7.48	4.24	3.15, 3.11	N1H 9.75; C2H 6.87; C4H 6.98; C5H 6.80; C6H 7.07; C7H 7.35
Ser 61	7.54	3.93	3.68	
Phe 62	7.86	5.25	2.42, 1.56	C2,6H 7.33; C3,5H 7.22; C4H 7.17
Tyr 63	8.22	5.50	3.02, 2.83	C2,6H 7.02; C3,5H 6.65
Leu 64	9.16	4.64	1.97, 1.78	γCH 1.78; δCH_3 1.03, 0.96
Leu 65	8.17	5.48	1.97, 1.54	γCH 1.56; δCH_3 1.03, 0.80
Tyr 66	9.13	5.36	3.04, 2.65	C2,6H 6.99; C3,5H 6.65
Tyr 67	8.94	5.98	3.21, 2.64	C2,6H 6.69; C3,5H 6.53
Thr 68	8.30	4.86	4.10	γCH_3 0.95
Glu 69	8.50	4.29	1.83, 1.71	γCH_2 1.87
Phe 70	8.76	4.83	2.78, 2.68	C2,6H 6.20; C3,5H 6.20; C4H, 5.72
Thr 71	8.22	4.46	3.85	γCH_3 0.85
Pro 72		4.57	2.40, 2.19	γCH 2.00, 1.40; δCH 3.97, 2.23
Thr 73	8.07	4.68	4.52	γCH_3 1.31
Glu 74	9.10	4.21	2.10, 2.06	γCH 2.33, 2.28
Lys 75	7.79	4.45	1.85, 1.77	γCH_2 1.38; δCH_2 1.65; ϵCH_2 2.98
Asp 76	7.13	5.12	2.79, 2.14	
Glu 77	8.62	4.81	2.06, 1.99	γCH_2 2.36

Table I (Continued)

residue	NH	αCH	βCH	γCH and others
Tyr 78	9.48	5.61	2.81, 2.71	C2,6H 7.10; C3,5H 6.89
Ala 79	8.80	5.05	1.19	
Cys 80	9.11	5.12	3.04, 2.61	
Arg 81	9.42	5.40	1.79, 1.18	CγH 1.32, 1.20; CδH 3.08, 2.90; εNH 7.09
Val 82	9.04	4.94	1.65	γCH_3 0.80, 0.60
Asn 83	9.03	5.19	2.80, 2.38	δNH 7.41, 6.62
His 84	7.74	4.58	2.88, 2.41	C2H 8.13; C4H 7.58; N3H 12.13
Val 85	8.06	4.00	1.98	γCH_3 0.86, 0.63
Thr 86	7.55	4.16	4.51	γCH_3 1.49; γOH 7.63
Leu 87	8.04	4.75	2.02, 1.81	γCH 1.71; δCH_3 0.95, 0.95
Ser 88	(8.94)	4.25	3.98, 3.96	
Gln 89	7.54	4.74	2.14, 1.88	γCH 2.28, 2.24; εNH 7.50, 6.81
Pro 90		4.50	1.84, 1.50	γCH 1.99, 1.77; δCH 3.73, 3.51
Lys 91	8.73	4.52	1.72, 1.64	γCH 1.32, 1.38
Ile 92	8.50	4.80	1.69	γCH 1.42; γCH_3 0.61; δCH_3 0.74
Val 93	9.06	4.32	1.88	γCH_3 1.03, 0.91
Lys 94	8.78	4.44	1.84, 1.84	γCH 1.53, 1.43
Trp 95	8.72	4.65	3.50, 2.61	N1H 10.51; C2H 7.12; C4H 7.99; C5H 7.46; C6H 6.97; C7H 7.62
Asp 96	8.46	4.49	2.76, 2.44	
Arg 97	7.46	3.42	1.42, 1.14	γCH 1.14, 0.92; δCH 2.95, 2.94; εNH (7.10)
Asp 98	8.20	4.65	2.77, 2.59	
Met 99	7.58	4.28	2.15, 2.01	γCH 2.56, 2.52; ϵCH_3 2.15

[a] Chemical shifts are in parts per million referenced to DSS and are accurate to ±0.02 ppm. Assignments of human β_2m at pH 5.9, 32 °C, are denoted in parentheses.

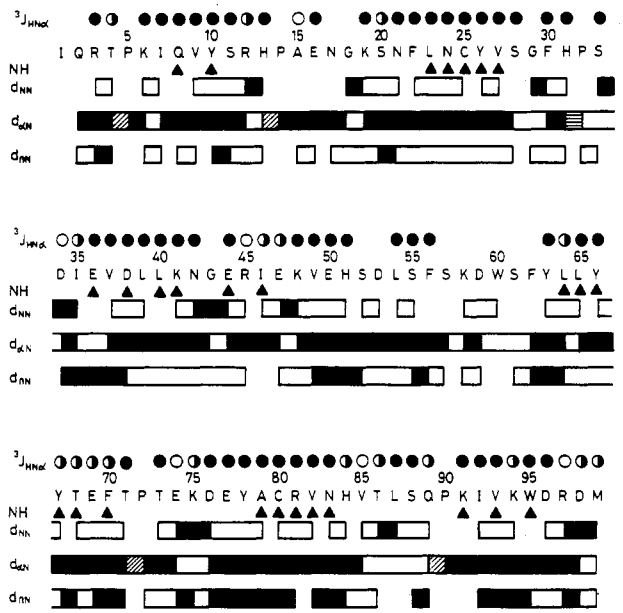

FIGURE 4: Primary sequence and summary of backbone coupling constants ($^3J_{HN\alpha}$), slow amide proton exchange rates (NH), and sequential NOE connectivities (d_{NN}, $d_{\alpha N}$, $d_{\beta N}$) observed for human β_2m. Filled, half-filled, and open circles indicate large (>8 Hz), intermediate (8 Hz > $^3J_{HN\alpha}$ > 6 Hz), and small (<6 Hz) coupling constants, respectively. Triangles indicate slow exchange rates (k_m < 10^{-2} min^{-1}). NOE connectivities are indicated by bars between successive residues (filled bars correspond to strong NOE's and empty bars to small NOE's). Sequential NOE's to the δCH_2 of proline are indicated by crosshatched bars. The sequential NOE $d_{\alpha\alpha}$ between His 31 and *cis*-Pro 32 is indicated by horizontally hatched bar.

missing at 400 MHz. Signals from N17, D53, L54, K58, and S88 in COSY or NOESY spectra were observed only at pH <6. Signals from Q2, G29, S57, and W60 were not found in COSY under any conditions, as for Q2 and S57 in NOESY. In these cases, αCH(i) was assigned using an NOE with NH(i+1), and NH(i) from NOE with αCH(i–1). A NH–αCH cross-peak from Phe 62 was observed as a weak signal only at a high temperature (47 °C).

A summary of sequential information obtained from NOESY spectra is given in Figure 4. NOE's were characterized as strong or weak on the basis of intensity. These NOE intensities basicly reflect changes for corresponding connectivities along the linear sequence.

All consecutive peaks for $d_{\alpha N}$ were found, ensuring continuity of sequential assignments. All d_{NN} were observed except for N21/F22, S28/G29, Y67/T68, and A79/C80 (see also Figure 5). A few others were not determined, due to small chemical shift differences (less than 0.2 ppm).

Table I lists all assignments obtained to date. A small number of additional resonances were observed, some only occasionally, presumably due to impurities or minor, alternative conformations of the protein.

Secondary Structure

A large number of intense sequential $d_{\alpha N}$ NOE's are evenly distributed along the backbone as seen in Figure 4, indicative of an extended backbone conformation. (The large number of strong $^3J_{HN\alpha}$ supports this conclusion.) The low number of intense d_{NN} and small $^3J_{HN\alpha}$ reflect a small number of loops and turns.

Direct evidence about backbone packing can be obtained from long-range backbone NOE connectivities: $d_{NN}(i,j)$, $d_{\alpha\alpha}(i,j)$ and $d_{\alpha N}(i,j)$ (Wüthrich, 1986). Long-range connectivities for d_{NN} and $d_{\alpha\alpha}$ were obtained from NOESY spectra shown in Figures 5 and 6. From these data, a structure pattern has been constructed and presented schematically in Figure 7. With a few exceptions, all connectivities for d_{NN} and $d_{\alpha\alpha}$ are consistent with this structure. The following properties of β_2m secondary structure are obvious:

(i) The large number of long-range $d_{\alpha\alpha}(i,j)$, $d_{\alpha N}(i,j)$, and $d_{NN}(i,j)$ NOE's confirm the existence of predominant β-structure.

(ii) More than 58 residues out of 99 are involved in forming two antiparallel β-sheets shown in Figure 7. An additional 5–7 residues do not belong to pure β-structure but can be attributed to this structure (for example bulges).

(iii) Both β-sheets are formed around a Cys residue (C25 or C80), being fixed in space through the disulfide bond with NOE's between the βCH_2's of Cys 25 and Cys 80.

Loops, Turns, and Bulges. Characteristics for loops and turns are intense sequential d_{NN}, weak sequential $d_{\alpha N}$, small $^3J_{HN\alpha}$ (<6 Hz), and the presence of Gly or Pro in the cor-

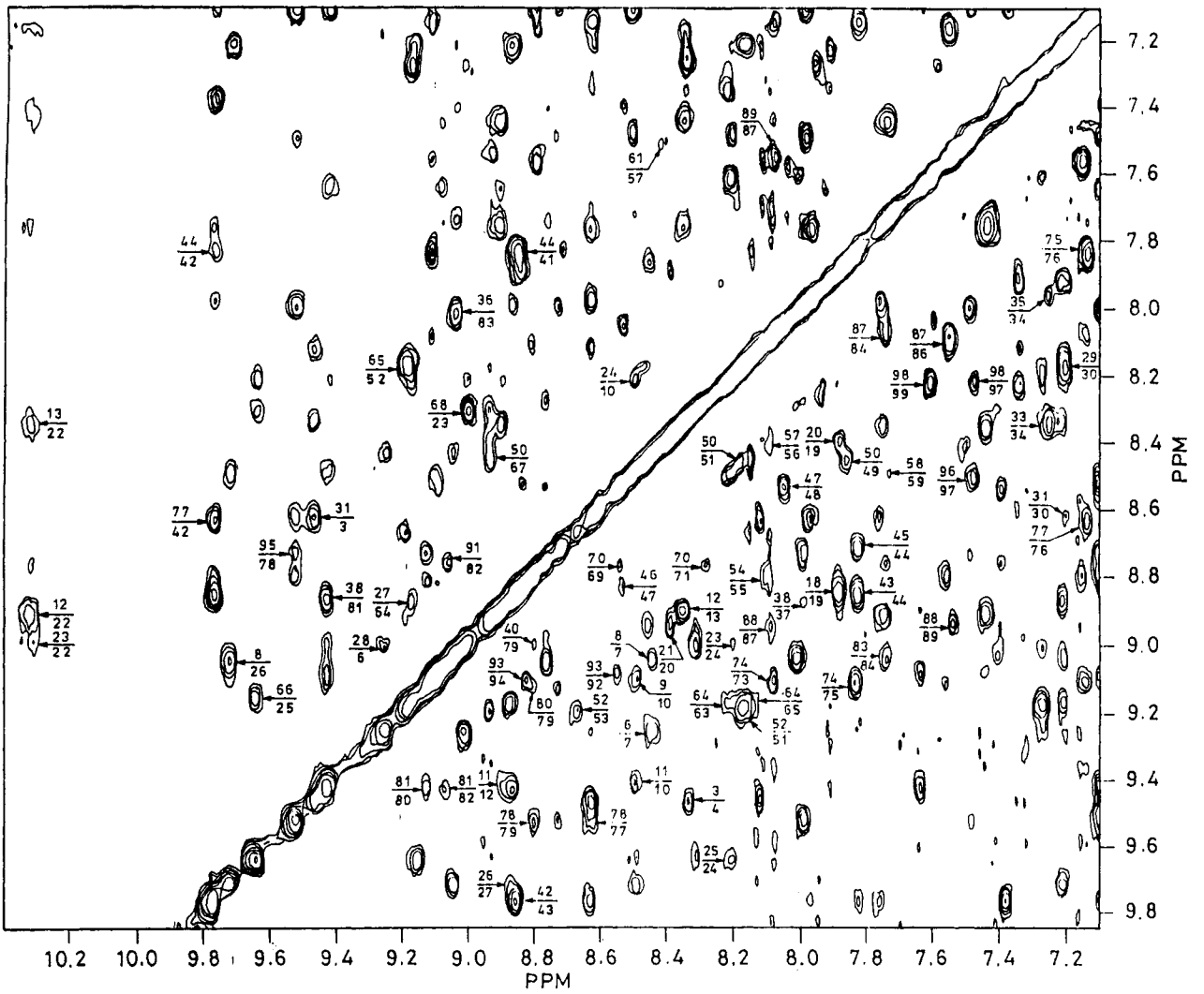

FIGURE 5: NH region of the phase-sensitive 600-MHz NOESY spectrum of human β_2m, τ_m = 200 ms. Conditions are the same as in Figure 2 but the temperature is 32 °C. Connectivities involving pairs of sequential NH protons (d_{NN}) are indicated mainly below the diagonal. Long-range d_{NN} connectivities are indicated mainly above the diagonal.

responding segment. The results indicate two relatively long loops for residues 13–19 and 72–76. From medium-range connectivities $d_{NN}(i,i+2)$ and $d_{\alpha N}(i,i+2)$ (Wagner et al., 1986), two tight turns (Richardson, 1981) were identified: 41–44 and 84–87.

Segment 41–44 with Gly in ideal $i+2$ position and visible $d_{NN}(42,44)$ could not be classified as any classical tight turn. Segment 84–87 was identified as a nearly classical type I tight turn with visible $d_{\alpha N}(85,87)$, moderate $d_{NN}(84,87)$, and small $^3J_{HN\alpha}$ for Val 85. Only $^3J_{HN\alpha}$ of Thr 86 is smaller than theoretical, being 7 Hz instead of 9 Hz.

The results summarized in Figures 4 and 7 indicate the presence of other turns or loops in local segments of the polypeptide chain: 32–35, 46–49, 57–61, and 87–90. However, the details of their conformation must await three-dimensional solution structure determination.

Another notable feature of the secondary structure is the identification of three β-bulges. β-Bulges have the effect of changing the register of strands in a β-sheet (Richardson et al., 1978). One β-bulge of classic or G1 type (G29, F30; F62) and two of wide type (T4, P5; F30 and D53, L54; L64) were found. The observation of two β-bulges next to each other in adjacent strands within a β-sheet is particularly interesting, i.e., bulges at residues 4–5 and 29–30 in strands I and II, respectively (Figure 7A).

Amide Exchange Rates. Twenty-six slowly exchanging amide protons were identified (Figures 4 and 7). Fourteen protons were kinetically evaluated with an upper limit of 10^{-2} min^{-1} for solvent-exchange rate constant k_m. It is no surprise that all slowly exchanging amide protons are embedded within the β-sheets. The ratio k_{mbond}/k_{mfree} can in some way be regarded as a measure of "packing density" and solvent accessibility. Taking Wüthrich's estimation of 10^2 min^{-1} for k_{mfree} (Wüthrich, 1986), a k_{mbond}/k_{mfree} of 10^{-4} characterizes β_2m as a moderately tight-packed structure.

Structural Inferences

Concentration and pH Effects. In the process of making assignments, experimental conditions such as concentration of β_2m, pH, and temperature of the solution were varied (see Materials and Methods). These variations caused small changes in resonance positions and facilitated the assignment process in crowded regions of the spectra. Spin connectivity maps remained almost identical, clearly indicating that no noticeable conformational changes occur over these experimental conditions with a few exceptions mentioned below.

Increasing the concentration of human β_2m higher than 10 mg/mL results in progressive line broadening to widths >10 Hz. This effect most proably is due to the tendency of β_2m to dimierze and form higher aggregates as reported for bovine

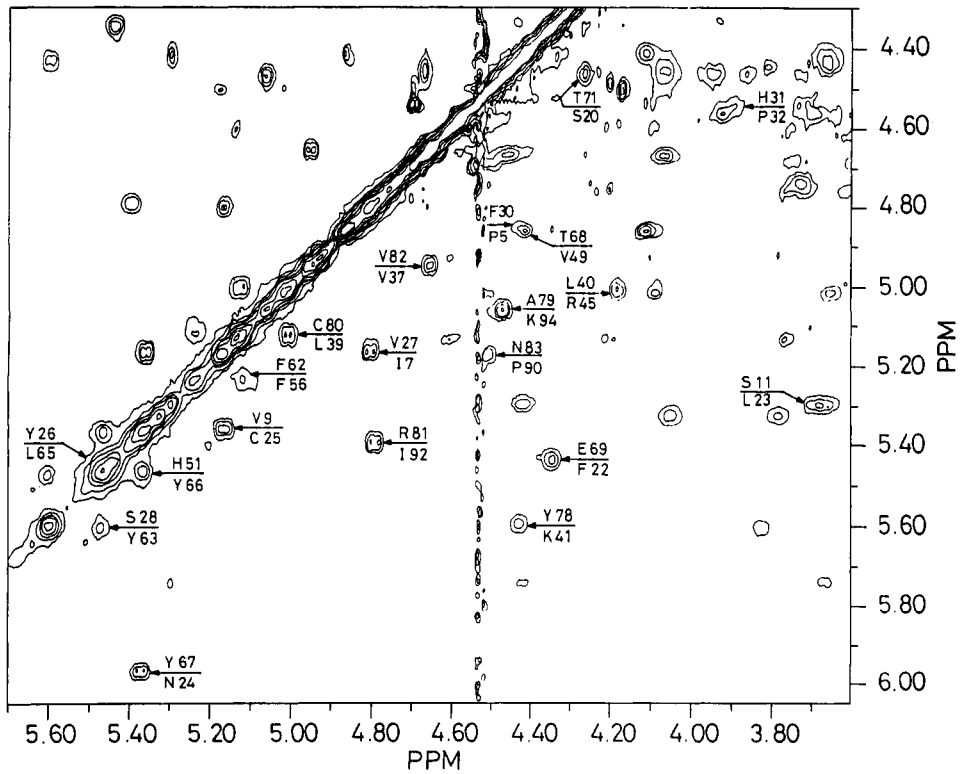

FIGURE 6: αCH region of the phase-sensitive 600-MHz NOESY spectrum of human β_2m in D_2O solution, τ_m = 200 ms at 8 mg/mL, pH* 6.94, 47 °C. Long-range and one sequential $d_{\alpha\alpha}$ (H31, P32) connectivities are shown.

Table II: Ring Current Effects and NOE's between Side Chains in Human β_2m

upfield-shifted signals	βCH_2	βCH_2	δCH_3	βCH	γCH_2	βCH	βCH	δCH	γCH_2
	L23	N24	I35	V37	K41	E50	S55	P72	R97
NOE with corresponding aromatic residues	F70		H84			Y66			
		Y10		Y66	Y78		Y63	F70	W95
	Y78		F30			Y67			

and guinea pig β_2m (Cigen et al., 1978; Kumosinski et al., 1981).

It was also found that human β_2m is most stable at pH 7 (physiological conditions). Either above or below pH 7, β_2m slowly denaturates. This could be related to the highly charged surface of this protein. Therefore, the assignments were mainly made at pH 7 and concentration below 10 mg/mL (Table I). Nevertheless, signals from NH of N17, D53, L54, K58, and S88 are missing from spectra at pH 7. Therefore, the H_2O spectra of β_2m recorded at pH 6 are shown in the figures of this paper and contain maximal information about individual connectivities.

Ring Current Shifts. A significant number of lines are strongly shifted upfield, more than 1 ppm, in comparison with "standard" positions in proteins (Grob & Kalbitzer, 1988). Such chemical shifts usually result from close proximity with aromatic rings (a ring current effect). NOE's were observed between these upfield-shifted signals and aromatic rings as summarized in Table II. As can be seen, besides interactions within the β-sheets, there is also evidence for close packing between the two sheets at three points: L23–Y78, V37–Y66, and I35–F30, indicating close side-chain interactions between the β-sheets.

Inaccessible Space. In the spectrum of β_2m in H_2O, signals from N3H of His 84 and from γOH of Thr 86 are observed, which is surprising due to expected fast exchange with solvent. In addition, C2H of His 84 does not shift with pH or exchange with D_2O as for the other three His residues. This effect is possible if His 84 and Thr 86 are protected from bulk H_2O. Probable protecting residues are Phe 30, Phe 62, and Ile 35 (see Table II). It is very likely that N1 of His 84 is hydrogen bonded with γOH of Thr 86 since NOE's are observed between C2H of His 84 and αCH, βCH, and γCH_3 of Thr 86. This protected space, especially His 84, probably plays an important role because in β_2m homologues His 84 is invariant while Thr 86 is substituted only in a single homologue, with Ser in mouse, with the γOH group being conserved.

Prolines. Direct proof for a cis peptide bond was obtained for Pro 32. An NOE exists between αCH and βCH_2 of His 31 and αCH of Pro 32. There are no NOE's between these protons and δCH_2 of Pro 32.

Protein Structural Flexibility. The behavior of the loop from Phe 56 to Phe 62 is very peculiar. There are two possibilities: it is either very rigid, or the rate of fluctuations between two or more structures are relatively slow on the NMR scale. On the basis of the following facts, the latter is more acceptable:

(i) All lines from NH and αCH protons in this loop, except residues 58 and 59, and some in proximity (Gly 29 and Pro 32) are significantly broader than any other lines in the spectrum.

(ii) At lower temperature, the signal from C2, 6H of Phe 56 progressively broadens and shifts upfield.

Solution and Crystal Structures of β_2m. In general, NMR results are in very good agreement with two recently reported X-ray secondary structures given in Table III (Bjorkman et al., 1987a,b; Garrett et al., 1989). It is clear that β_2m is a pure β-sheet protein. Nearly all β-sheet residues found from the X-ray crystal structure in the solid stay within β-sheet structures in solution as determined from NMR (Figure 7).

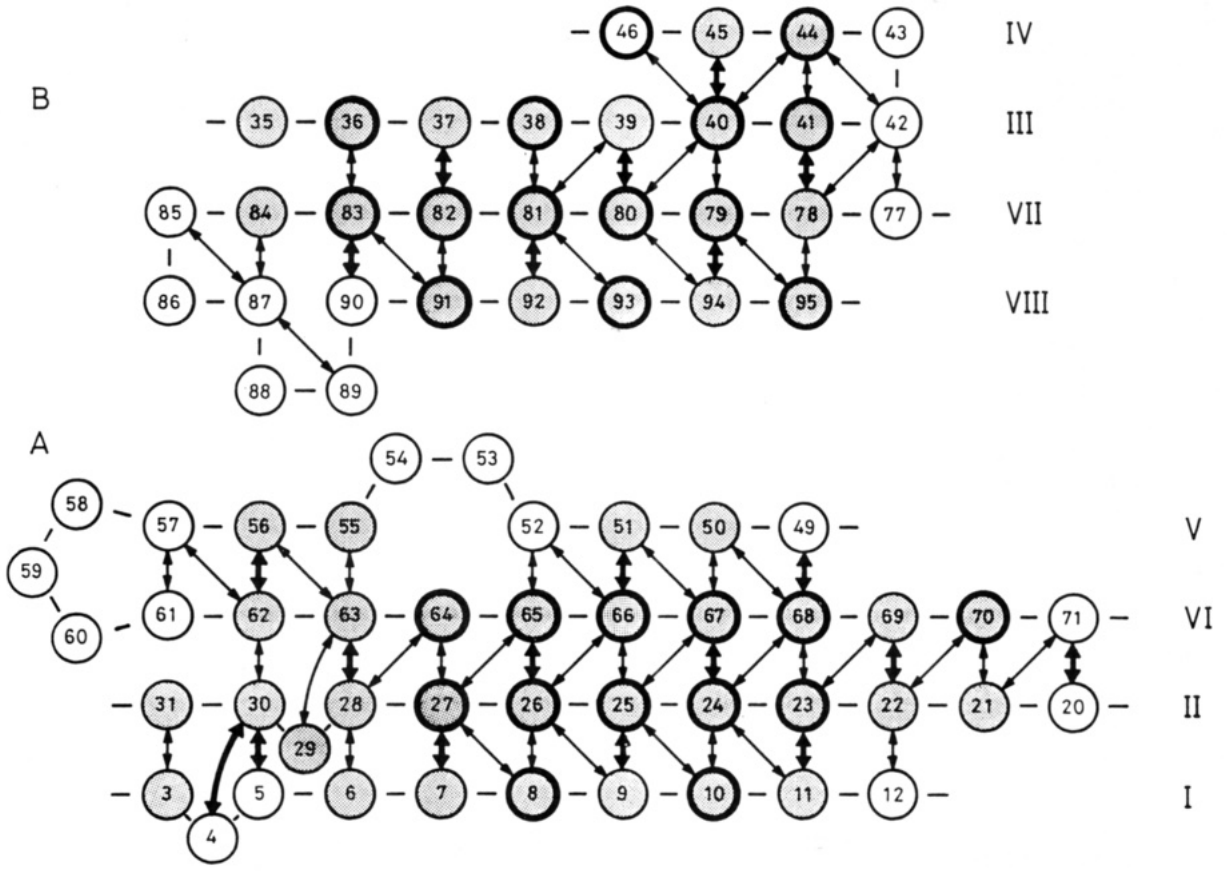

FIGURE 7: Schematic representation of the secondary structure of human β_2m as determined by analysis of NOE connectivities and amide NH exchange. Observed middle- and long-range NOE's are indicated by arrows [$d_{\alpha\alpha}(i,j)$ by thick and $d_{NN}(i,j)$ and $d_{\alpha N}(i,j)$ by thin arrows]. Residues with slow amide NH exchange rates are denoted by thickened circles. The β-strands are labeled with roman numerals in the order in the primary sequence. Residues included in β-structure according to X-ray data are indicated by stippled circles.

Table III: Comparison of the Secondary Structure of Human β_2m from NMR Analysis (this paper) and from X-ray Analysis for Histocompatibility Antigens HLA–A2 and HLA–Aw68 Obtained from the Brookhaven Protein Data Bank

X-ray (Brookhaven Protein Data Bank)		
HLA-A2.1 strand	HLA-Aw68.1 strand	NMR (this paper) strand
	β-sheet I	
6–11	3, 6–11	3, 6–12
21–30	21–31	20–28, 31
62–70	62–70	61–71
50–51, 55–56	50–51, 55–56	50–52, 55–57
	β-sheet II	
44–45	44–45	44–45
35–41	36–41	36–41
78–84	78–83	77–83
91–94	91–95	90–95
	turns	
	16–19	?
	41–44	41–44
	57–60	57–61
		84–87

However, some differences are observed. Arg 3 really does belong to β-structure as published for HLA-Aw68.1. Strand 21–31 is not so simple. First, it extends to Ser 20. Second, residues 29 and 30 can be regarded as some kind of "bulge" irregularity rather than pure β-structure. In this context, Thr 4 and Pro 5 are in contact with Phe 30 and can be attributed to another bulge. A tight turn for Glu 16–Lys 19 connecting β-strands 6–11 and 21–31 was found in the X-ray crystal structure. This region has been assigned as a loop in solution because of insufficient information with only a strong d_{NN} between Gly 18 and Lys 19 but no d_{NN} between Glu 16 and Lys 19 (Figure 5).

Differences appear even smaller for the second β-sheet. It was impossible to determine if Ile 35 and His 84 belong to β-strands since their αCH chemical shifts are identical. In contrast with X-ray data, a tight turn for His 84–Leu 87 was found.

It seems that differences in secondary structure between crystal and solution are due to single residue variations in defining β-strand termini. Thus, β-strands 20–31, 77–83, and 90–95 defined by NMR have a single extra residue at their N-terminus compared with X-ray results; NMR determined β-strand 6–12 has an extra residue at its C-terminus; and NMR β-strands 61–71 and 49–52 have one extra residue at each end. Ser 57 is included in strand 55–57 in solution. Thus, the current NMR secondary solution structure is essentially the same as the crystal structure of human β_2m. Single-residue discrepancies in defining β-strand termini might relate to how stringent β-strand criteria are for NMR and X-ray data. It should be emphasized that the same structure is found in spite of the fact that monomeric human β_2m is used in NMR studies whereas human class I histocompatibility antigens, i.e., dimeric complexes, were used in crystal studies.

REFERENCES

Bax, A., & Drobny, G. (1985) *J. Magn. Reson. 61*, 306–320.

Becker, J. W., & Reeke, G. N., Jr. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4225–4229.

Berggard, I., & Bearn, A. G. (1968) *J. Biol. Chem. 243*, 4095–4103.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.
Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.
Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., & Wiley, D. C. (1987a) *Nature 329*, 506–512.
Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., & Wiley, D. C. (1987b) *Nature 329*, 512–518.
Bodenhausen, G., Koglen, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370–388.
Chothia, C. (1973) *J. Mol. Biol. 75*, 295–302.
Chothia, C., & Janin, J. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4146–4150.
Cigen, R., Ziffer, J. A., Berggard, B., Cunningham, B. A., & Berggard, I. (1978) *Biochemistry 17*, 947–955.
Cunningham, B. A., Wang, J. L., Berggard, I., & Peterson, P. A. (1973) *Biochemistry 12*, 4811–4822.
Garrett, T. P. J., Saper, M. A. Bjorkman, P. J., Strominger, J. L., & Wiley, D. C. (1989) *Nature 342*, 692–696.
Gejyo, F., Yamada, T., Odani, S., Nakagawa, Y., Arakawa, M., Kunitomo, T., Kataoka, H., Suzuki, M., Hirasawa, Y., Shirahama, T., Cohen, A. S., & Schmid, K. (1985) *Biochem. Biophys. Res. Commun. 129*, 701–706.
Gorevic, P. D., Casey, T. T., Stone, W. J., DiRaimondo, C. R., Prelli, F. C., & Frangione, B. (1985) *J. Clin. Invest. 76*, 2425–2429.
Grey, H. M., Kubo, R. T., Colon, S. M., Poulik, M. D., Cresswell, P., Springer, T., Turner, M., & Strominger, J. L. (1973) *J. Exp. Med. 138*, 1608–1612.
Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870–7872.
Grob, K. H., & Kalbitzer, H. R. (1988) *J. Magn. Reson. 76*, 87–99.
Gurevich, A. Z., Barsukov, I. L., Arseniev, A. S., & Bystrov, V. F. (1984) *J. Magn. Reson. 56*, 471–478.
Hall, P. W., & Vasiljević, M. (1973) *J. Lab. Clin. Med. 81*, 897–904.
Hore, P. J. (1983) *J. Magn. Reson. 54*, 539–542.
Karlsson, F. A., & Lenkei, R. (1977) *Scand. J. Clin. Lab. Invest. 37*, 169–173.
Kumar, A., Wagner, G., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 96*, 1156–1163.
Kumosinski, T. F., Brown, E. M., & Groves, M. L. (1981) *J. Biol. Chem. 256*, 10949–10953.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Neuhaus, D., Wagner, G., Vasak, M., Kagi, J. H. R., & Wüthrich, K. (1985) *Engl. J. Biochem. 151*, 257–273.
Peterson, P. A., Evrin, P. E., & Berggard, I. (1969) *J. Clin. Invest. 48*, 1189–1198.
Peterson, P. A., Rask, L., & Lindblom, J. B. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 35–39.
Ploegh, H. L., Orr, H. T., & Strominger, J. L. (1981a) *Cell 24*, 287–299.
Ploegh, H. L., Orr, H. T., & Strominger, J. L. (1981b) *J. Immunol. 126*, 270–275.
Pras, M., Prelli, F., Franklin, E. C., & Frangione, B. (1983) *Proc Natl. Acad. Sci. U.S.A. 80*, 539–542.
Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.
Richardson, J. S., Getzoff, E. D., & Richardson, D. C. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2574–2578.
Salemme, F. R. (1983) *Prog. Biophys. Mol. Biol. 42*, 95–133.
Salemme, F. R., & Weatherford, D. W. (1981) *J. Mol. Biol. 146*, 119–141.
Skinner, M., & Cohen, A. S. (1981) *Biochem. Biophys. Res. Commun. 99*, 1326–1332.
Smithies, O., & Poulik, M. D. (1972) *Science 175*, 187–189.
Trollfors, B., & Norrby, R. (1981) *Nephron 28*, 196–199.
Tse, D. B., & Pernis, B. (1984) *J. Exp. Med. 159*, 193–207.
Tveteraas, T., Sletten, K., & Westermark, P. (1985) *Biochem. J. 232*, 183–190.
Vicent, C., Revillard, J. P., Galland, M., & Traeger, J. (1978) *Nephron 21*, 260–268.
Wagner, G. (1983) *J. Magn. Reson. 55*, 151–156.
Wagner, G., Neuhaus, D., Worgotter, E., Vašok, M., Kagi, J., & Wüthrich, K. (1986) *J. Mol. Biol. 187*, 131–135.
Wibell, L., Evrin, P. E., & Berggard, I. (1973) *Nephron 10*, 320–331.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.