Correction in the reported amino acid sequence for human seminal plasma prostatic inhibin

Vinit K. Rastogi, K. V. R. Chary and Girjesh Govil

Chemical Physics Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India

NMR studies on human seminal plasma prostatic inhibin (HSPI) reveal that the amino acid at position 93 is Ile and not Gly as was reported earlier. This suggests that HSPI and prostatic secretory protein are one and the same protein.

HUMAN seminal plasma contains a variety of proteins secreted by seminal vesicles and the prostate. Human seminal plasma prostatic inhibin (HSPI), prostatic secretory protein (PSP-94), immunoglobulin binding factor (IgBF) and β-microseminoprotein (β-MSP) are among the human seminal proteins which have been isolated, purified and characterized during the last decade. All these proteins have close (~99%) sequence homology. HSPI, first isolated by Sheth and his coworkers[1], has biological activity ranging from preventing pregnancy to curing prostate cancer. Inhibin prevents pregnancy in mammals by modulating the level of circulating follicle stimulating hormones (FSH)[2]. It suppresses prolactin, a hormone that promotes lactation[3]. Therefore, neutralizing inhibin through active immunization leads to increase in milk production. Seidah *et al.* have reported the primary sequence of HSPI[4]. This was followed by the discovery of PSP-94 (ref. 5) which showed complete sequence homology with that of HSPI except for one amino acid residue at position 93. HSPI is reported to contain Gly at this site whereas PSP-94 has Ile (Figure 1).

Around the same time, Kazuko *et al.* isolated β-MSP[6] which was reported to be a 93 residue protein showing differences in sequence homology at positions 39 and 40 from PSP-94. While β-MSP contains Pro and Thr at these positions, PSP-94 and HSPI have Thr and Cys, respectively (Figure 1). Another group, Zhi *et al.*[7] who purified and characterized IgBF, reported the primary sequence of only thirty N-terminus and two C-terminal amino acids of IgBF, which show complete sequence homology with the corresponding stretches of PSP-94. On the basis of such sequence homologies, Zhi *et al.*[7] suggested that HSPI, PSP-94, β-MSP and IgBF may be the same protein.

No information is available about the three-dimensional (3D) structure of these proteins. We have undertaken studies on the 3D structure of HSPI using multidimensional NMR spectroscopy. Our investigations unequivocally establish that the amino acid at position 93 is Ile and not Gly.

Semen samples of healthy men obtained at infertility clinics were pooled. The material was centrifuged at 3000 rpm for 10 min at 4°C. The pellet containing sperms was discarded. The supernatant was mixed with four volumes of freshly distilled chilled ethanol, allowed to settle overnight at 4°C, and then centrifuged at 3000 rpm for 30 min to separate precipitated proteins. The precipitate was washed with chilled acetone followed by chilled ether, and finally air dried to yield the crude extract which served as the starting material for purification of the protein.

The crude extract powder was dissolved in 100 ml 0.05 M sodium deuteroacetate buffer, pH 4.0. Deuterated buffer was used to avoid the interference of resonances arising from the residual acetate in the NMR spectrum. Fractionation of the crude extract was effected on a Sephadex G-100 column (110 cm × 5 cm) (Pharmacia) using 0.05 M sodium acetate, pH 4.0 for equilibration and elution. Fractions monitored at 280 nm were collected at a flow rate of 1 ml/min and pooled on the basis of the elution profile. The active fraction from sephadex G-100 chromatography was concentrated by lyophilization and equilibrated by dialysis with tris-HCl buffer (0.05 M), pH 8.0. About 1 g of immunoactive material was loaded on a DEAE cellulose column (30 cm × 2.5 cm). The column was first washed with tris buffer to remove the unadsorbed proteins till it showed negligible OD at 280 nm using a flow rate of 0.5 ml/min. The adsorbed proteins were then eluted with Tris buffer containing 0.0–0.2 M sodium chloride continuous gradient. The immunoactive fractions were further purified by HPLC on a gel permeation preparative column (Ultropack TSK G-2000 SWG; 21.5 mm × 600 mm) with an exclusion limit of 200,000 daltons. The proteins were eluted with 0.05 M acetate buffer, pH 4.0 at a flow rate of 3 ml/min. The HPLC purified material was dialysed and lyophilized. The activity of HSPI was checked by radioimmunoassay[8].

Two-dimensional (2D) NMR experiments were carried out on a Varian Unity plus spectrometer with a ^1H frequency of 600 MHz. The 2D experiments in ^2H_2O included two-quantum-filtered correlation spectroscopy (2QF-COSY)[9], three-quantum-filtered correlation spectroscopy (3QF-COSY)[10], clean total correlation spectroscopy (clean-TOCSY)[11] with a mixing time of 80 ms, and nuclear Overhauser enhancement spectroscopy (NOESY)[12] with a mixing time of 150 ms. Spectra in 90% H_2O + 10% ^2H_2O included WATERGATE NOESY[13] with a mixing time of 200 ms and clean-TOCSY with a mixing time of 100 ms.

In the previously reported HSPI sequence of 94 residues all amino acid residues except alanine are present. The number of various amino acids is as follows: Gly: 5, Val: 6, Leu: 2, Ile: 6, Asp: 5, Asn: 5, Cys: 10, Ser: 7, Thr: 8, Glu: 8, Gln: 2, Met: 1, Arg: 2, Lys: 11, Pro: 5, Phe: 2, Tyr: 4, His: 1, Trp: 2. HSPI is particularly

rich in lysines and cysteines. Though the occurrence of 11 lysine residues in a protein of 94 amino acids is not uncommon, the number of cysteine residues[10] is much higher than the statistical average for proteins of similar size. The spin systems of most of these residues have been identified using 2D NMR spectra recorded in ^2H_2O solution, which has been followed by sequence-specific resonance assignments using standard strategies[14]. The procedures involve identifying individual amino-acid spin systems using J-correlated spectroscopy followed by detection of sequential cross peaks arising from short distances (d_{NN}, $d_{\alpha N}$, $d_{\beta N}$ etc.) in the NOESY spectrum[14]. A relatively large percentage of unique dipeptide segments in HSPI, makes sequence-specific ^{1}H-NMR assignments much simpler. Search for a possible ambiguity in the primary sequence at position 93 started from the sequence-specific assignment of the two tryptophans (Trp32 and Trp92). In the NOESY spectrum Trp N(ε^1)H proton is easily identifiable because of its inherent chemical shift which is most down-field shifted (~ 10.3 ppm) in the ^{1}H spectrum (Figure 2). Two distinct nOes are observed for Trp N(ε^1)H arising from C(δ^1)H and C(ζ^2)H of the indole rings. The distinction between the two protons has been achieved from the observation of N(ε^1)H–C(δ^1)H cross peak in the TOCSY spectrum in H_2O (not shown here). N(ε^1)H and C(ζ^2)H are not J-coupled hence do not show mutual interaction in the TOCSY spectrum. However, since the inter-proton distance is 2.84 Å one observes strong nOe between these protons which helps in the identification of C(ζ^2)H (Figure 2). From the knowledge of C(ζ^2)H position,

	1			5		32										42			92		94	
HSPI	S	C	Y	F	I	—	W	Q	T	D	N	C	E	T	C	T	C	—	W	G	I	COOH
PSP-94	S	C	Y	F	I	—	W	Q	T	D	N	C	E	T	C	T	C	—	W	I	I	COOH
β-MSP	S	C	Y	F	I	—	W	Q	T	D	N	C	E	P	T	T	C	—	W	I		COOH
IgBF	S	C	Y	F	I	—														I	I	COOH

Figure 1. Amino acid sequence of HSPI and related proteins isolated from human seminal plasma.

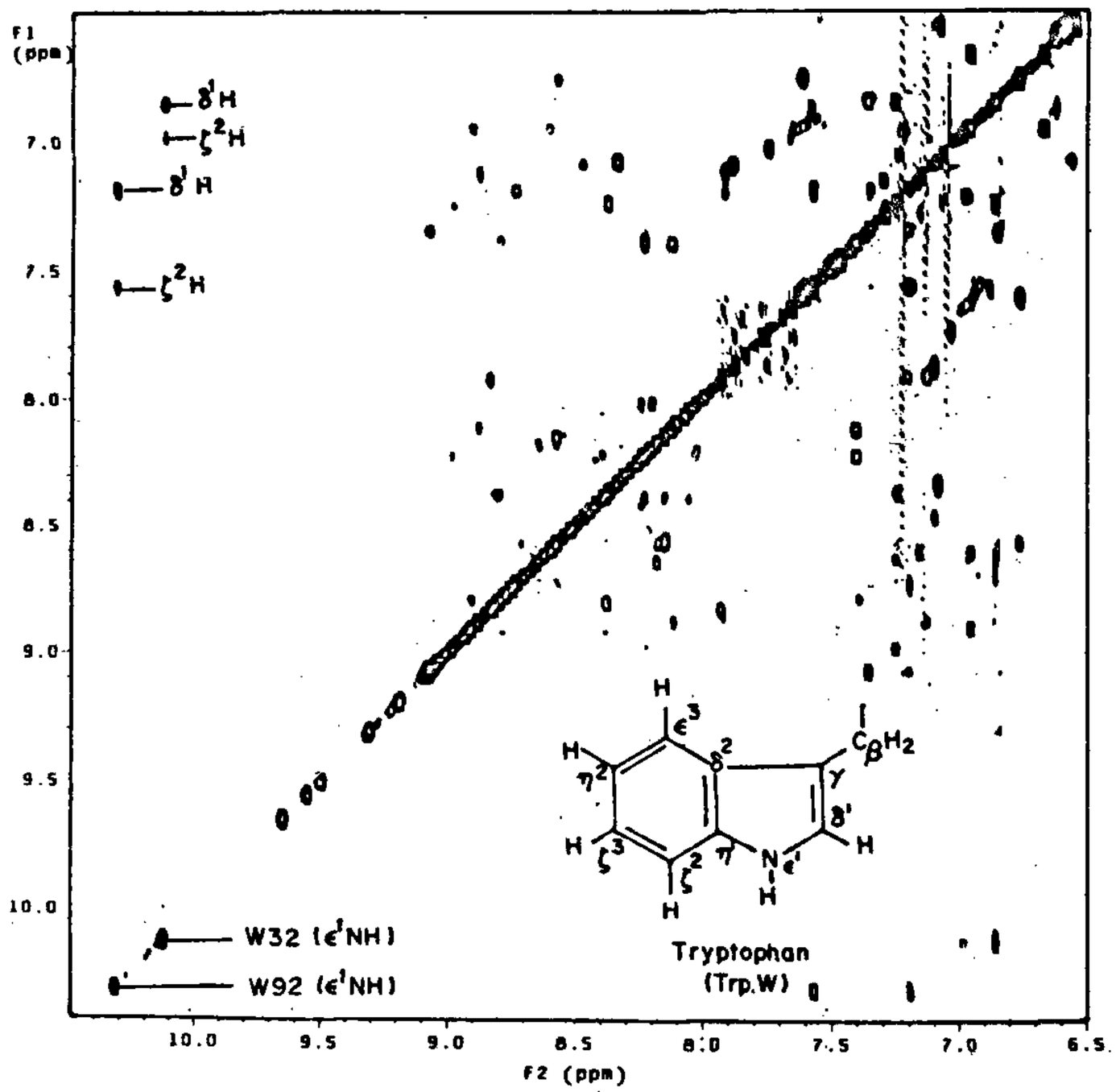

Figure 2. Selected region of the water-gated NOESY spectrum of HSPI recorded in 90% H_2O + 10% ^2H_2O at 20°C and pH 3.0. This shows the nOe connectivities from the side chain N(ε^1)H of two tryptophans to their respective C(δ^1)H and C(ζ^2)H protons. Experimental parameters: $t_1$max = 37.5 ms; $t_2$max = 256 ms; recycle delay = 1 s, 160 scans/t_1 increment. TD = 600 and 4096 along t_1 and t_2 dimensions, respectively. The data were multiplied with sine bell window functions shifted by $\pi/4$ and $\pi/8$ along t_1 and t_2 axes, respectively and zero filled to 1024 data points along t_1 dimension prior to 2D-FT. Digital resolution along ω_1 and ω_2 corresponds to 7.8 and 3.9 Hz/pt, respectively. The inset shows the indole ring of tryptophan with the standard nomenclature used for numbering the hydrogen atoms.

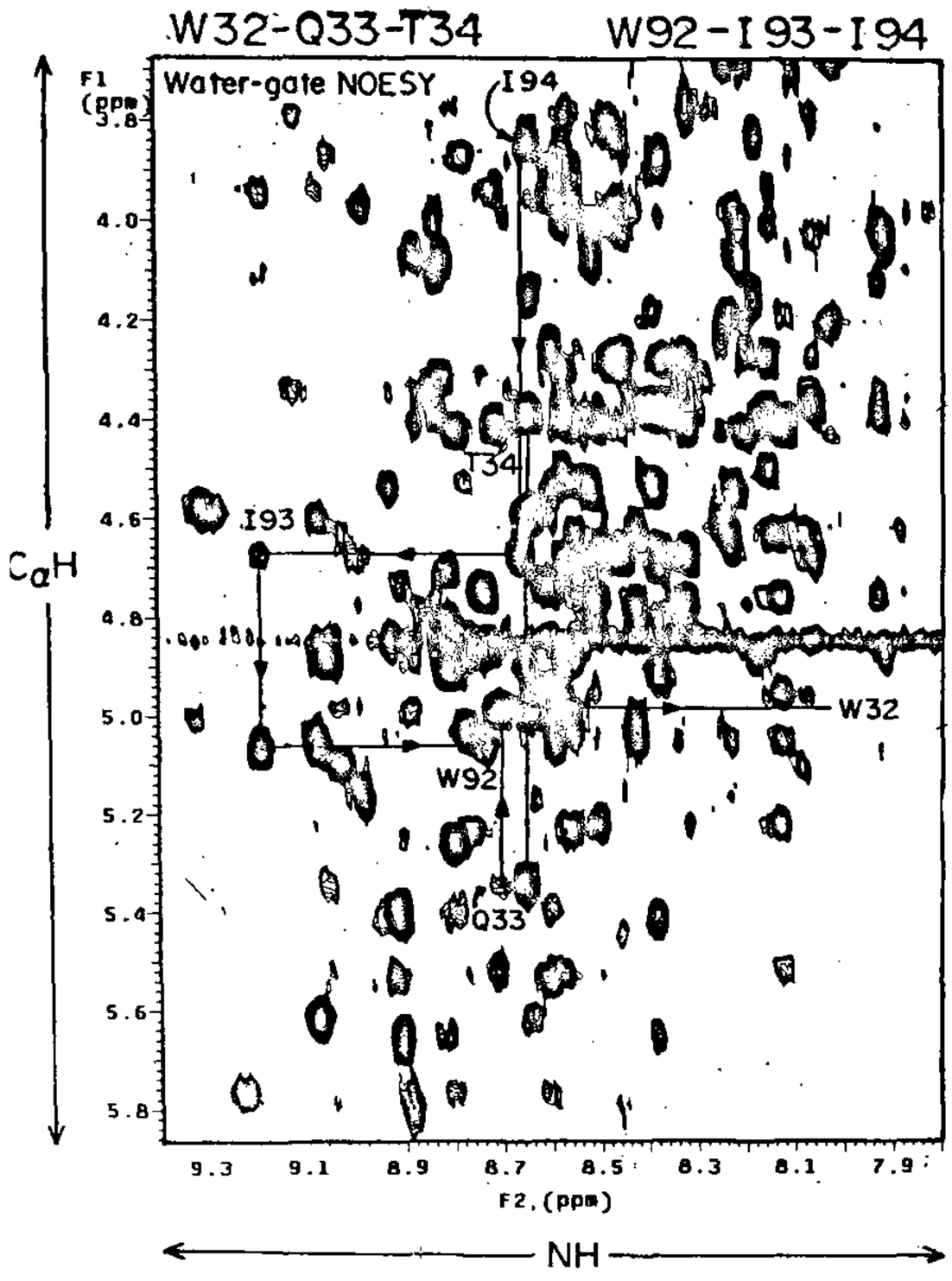

Figure 3. Finger-print region of the WATERGATE NOESY showing sequential connectivities (solid lines) for the protein segment Trp32–Gln33–Thr34 and Trp92–Ile93–Ile94.

the rest of the indole ring protons C(ζ^3)H, C(η^1)H and C(ε^2)H have been assigned from the TOCSY spectrum in ^2H_2O (not shown here). We could identify both sets of indole proton resonances arising from the two Trp residues. Further, the intra-residue nOes C(δ^1)H to C(α)H and C(β)H protons and C(α)H to NH (backbone) allowed complete identification of the remaining protons of the tryptophans. With this knowledge, nOe cross-peaks corresponding to sequential distances $d_{\alpha N}$, $d_{\beta N}$ and d_{NN} between neighbouring amino acid residues were assigned. This led to assignment of the glutamine and isoleucine spin systems (Figure 3), as sequential neighbours of Trp32 and Trp92, respectively. Sequential connectivities of Gln33 and Ile93 with Thr34 and Ile94, respectively confirm these assignments (Figure 3). The fact that segments Trp32–Gln33–Thr34 and Trp92–Ile93–Ile94 are unique in the polypeptide primary sequence, substantiates the sequence-specific resonance assignment. Thus, it could be unequivocally established that the amino acid at position 93 is Ile and not Gly as reported earlier. This is further supported by the identification of only four Gly residues in the COSY spectrum instead of five expected, if the amino acid at position 93 was Gly. Such a sequence correction suggests that HSPI and PSP-94 are one and the same protein.

1. Thakur, A. N., Vaze, A. Y., Dattreyamurthy, B. and Sheth, A. R., *Indian J. Exp. Biol.*, 1981, **19**, 307–313.
2. Garde, S., Sheth, A. R., Lohiya, N. K. and Shah, M. G., *J. Biosci.*, 1992, **17**, 67–85.
3. Moodbidri, S. B., Hurkadli, K. S. and Sheth, A. R., *Indian J. Exp. Biol.*, 1989, **27**, 214–216.
4. Seidah, N. G., Arbatti, N. J., Rochemont, J., Sheth, A. R. and Chretian, M., *FEBS Lett.*, 1984, **175**, 349–355.
5. Mbikay, M., Nolet, S., Fournier, S. *et al.*, *DNA*, 1987, **6**, 23–29.
6. Akiyama, K., Yoshioka, Y., Schmid, K., Offner, G. D., Troxler, R. F., Tsuda, R. and Hara, M., *Biochim. Biophys. Acta*, 1985, **829**, 288–294.
7. Liang, Z. G., Kamada, M. and Koide, S. S., *Biochem. Biophys. Res. Commun.*, 1991, **1**, 356–359.
8. Vaze, A. Y., Thakur, A. N. and Sheth, A. R., *J. Reprod. Fert. Suppl.*, 1979, **26**, 135–146.
9. Rance, M., Sorenson, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. Wuthrich, K., *Biochem. Biophys. Res. Commun.*, 1983, **117**, 479–485.
10. Muller, N., Ernst, R. R. and Wuthrich, K., *J. Am. Chem. Soc.*, 1986, **108**, 6482–6492.
11. Griesinger, C., Otting, G., Wuthrich, K. and Ernst, R. R., *J. Am. Chem. Soc.*, 1988, **110**, 7870–7872.
12. Anil, K., Wagner, G., Ernst, R. R. and Wuthrich, K., *Biochem. Biophys. Res. Commun.*, 1980, **96**, 1156–1163.
13. Piotto, M., Saudek, V. and Sklenar, V., *J. Biomol. NMR*, 1992, **2**, 661–665.
14. Wuthrich, K., *NMR of Proteins and Nucleic Acids*, Wiley, New York, 1986.

ACKNOWLEDGEMENTS. We thank Late Dr A. R. Sheth, Dr S. B. Moodbidri and Mr M. S. Kadam for their help provided during the course of isolation and purification of HSPI. The facilities provided by the National Facility for High Field NMR, Tata Institute of Fundamental Research, Mumbai are gratefully acknowledged.

Received 7 November 1996; accepted 25 December 1996