doi:10.1016/S0022-2836(02)01471-7 *J. Mol. Biol.* (2003) **326**, 1549–1557

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

NMR Structure of the Human Doppel Protein

Thorsten Lührs*, Roland Riek, Peter Güntert and Kurt Wüthrich*

Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule Zürich CH-8093 Zürich, Switzerland

**Corresponding authors*

The NMR structure of the recombinant human doppel protein, hDpl(24–152), contains a flexibly disordered "tail" comprising residues 24–51, and a globular domain extending from residues 52 to 149 for which a detailed structure was obtained. The globular domain contains four α-helices comprising residues 72–80 (α1), 101–115 (α2^a), 117–121 (α2^b), and 127–141 (α3), and a short two-stranded anti-parallel β-sheet comprising residues 58–60 (β1) and 88–90 (β2). The fold of the hDpl globular domain thus coincides nearly identically with the structure of the murine Dpl protein. There are close similarities with the human prion protein (hPrP) but, similar to the situation with the corresponding murine proteins, hDpl shows marked local differences when compared to hPrP: the β-sheet is flipped by 180° with respect to the molecular scaffold formed by the four helices, and the β1-strand is shifted by two residues toward the C terminus. A large solvent-accessible hydrophobic cleft is formed on the protein surface between β2 and α3, which has no counterpart in hPrP. The helix α2 of hPrP is replaced by two shorter helices, α2^a and α2^b. The helix α3 is shortened by more than two turns when compared with α3 of hPrP, which is enforced by the positioning of the second disulfide bond in hDpl. The C-terminal peptide segment 144–149 folds back onto the loop connecting β2 and α2. All but four of the 20 conserved residues in the globular domains of hPrP and hDpl appear to have a structural role in maintaining a PrP-type fold. The conservation of R76, E96, N110 and R134 in hDpl, corresponding to R148, E168, N183 and R208 in hPrP suggests that these amino acid residues might have essential roles in the so far unknown functions of PrP and Dpl in healthy organisms.

Keywords: doppel protein; prion protein; transmissible spongiform encephalopathy; NMR structure

Introduction

The prion protein (PrP) is a host-encoded protein,[1,2] which is ubiquitous in healthy mammalian organisms in the benign cellular form, PrP^C, and which has been implicated to have a role in transmissible spongiform encephalopathies (TSEs).[3] In particular, prion protein knockout mice of the strain *ZHI-Prnp*$^{0/0}$ were found not to be susceptible to prion infection,[4] which established a stringent requirement for the presence of host-encoded PrP^C for the onset of a TSE or "prion disease". This was confirmed by subsequent reintroduction of PrP-genes that code either for full-length PrP or for various N-terminally truncated variants thereof, which re-established the susceptibility of the resulting mouse strains to TSE-infection.[5,6] While the *ZHI-Prnp*$^{0/0}$ mice and another line of *Prnp* knockout mice[7] showed no obvious phenotype, two other strains of *Prnp* knockout mice, *Ngsk-Prnp*$^{0/0}$ and *Rcm0-Prnp*$^{0/0}$,

Present addresses: T. Lührs and R. Riek, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA; P. Güntert, RIKEN Genomic Sciences Center, W505, 1-7-22 Suehiro, Tsurumi, Yokohama 230-0045, Japan.

Abbreviations used: NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; ^{15}N{^{1}H}-NOE, heteronuclear Overhauser enhancement of ^{15}N after saturation of ^{1}H; rmsd, root-mean-square deviation; PrP, prion protein; PrP^C, cellular form of PrP; PrP^{Sc}, scrapie form of PrP; Dpl, doppel protein; hDpl(24–152), fragment of human Dpl comprising residues 24–152; mDpl(26–157), fragment of murine Dpl comprising residues 26–157; hPrP, human prion protein; TSE, transmissible spongiform encephalopathy; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight.

E-mail addresses of the corresponding authors: luehrs@salk.edu; wuthrich@mol.biol.ethz.ch

developed progressive signs of ataxia within 70 weeks after birth.[8–10] This apparent discrepancy of the observations with different knockout mouse strains was rationalized only recently through the discovery of a novel gene locus (*Prnd*) 16 kb downstream of *Prnp* and its product, the doppel protein (Dpl).[11]

The natural functions of Dpl as well as PrP^C are unknown. It has been argued that either an involvement of Dpl in prion diseases or a role in neural differentiation is unlikely.[12] In this situation, it is an intriguing observation that two neurologic disorders presenting in concert with overexpression of two distinct proteins appear both to be cured by the additional expression of wild-type PrP^C. The sequence identity between Dpl and PrP is only about 20%, so that experimental determination of Dpl 3D-structures is a significant addition to the data available as a foundation for functional studies of these proteins. So far, in addition to the NMR structures of recombinant murine PrP,[13,14] Syrian hamster PrP,[15,16] human PrP[17] and bovine PrP,[18] the structure of murine Dpl[19] has been published. Here, we describe the NMR structure determination of hDpl, and compare this new structure with mDpl and the aforementioned wild-type prion proteins.

Results

Resonance assignment

Sequence-specific assignments for the backbone and side-chain atoms were obtained using triple resonance, HCCH-total correlated spectroscopy (TOCSY) and [^{1}H,^{1}H]-nuclear Overhauser enhancement spectroscopy (NOESY) experiments with a ^{13}C,^{15}N-labeled sample of hDpl(24–152) (for details, see Materials and Methods). The assignments are complete except for the backbone amide protons of Y91, K143, C145 and F147, all side-chain protons of R32, K34 and K143, H^{δ} of R27, $H^{\varepsilon 1}$ of H31, H^{γ} of P86, H^{β} of D87, H^{β} of I89, H^{β} of C145 and H^{ζ} of F147. For all proline residues except P86, $^{13}C^{\beta}$ chemical shifts near 32 ppm were observed, indicating that they are in the *trans* conformation.[20] For all Xxx-Pro bonds, including the dipeptide segment 85–86, the *trans* conformation was confirmed independently by strong sequential $d_{\alpha\delta}$ or $d_{\beta\delta}$ NOE connectivities.[21]

In the [^{15}N,^{1}H]-correlated spectroscopy (COSY) spectra, resonance doubling was observed for some of the residues. The same resonance doublings were seen in three separately prepared batches of the protein. Both SDS-PAGE and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectra showed that the protein was homogeneous to at least 95%, ruling out the possibility that the resonance doublings with approximately 1:1 intensities were caused by a contaminating protein. The oxidization state of the four Cys residues was checked by incubating the protein with iodacetamide in 6 M guanidinium hydrochloride (GdHCl). Only if the protein was reduced with dithiothreitol prior to the incubation, four acetamide moieties were incorporated, as measured by MALDI-TOF mass spectroscopy. Otherwise, the protein was inert towards treatment with iodoacetamide, demonstrating that two disulfide bonds were formed. The resonance doublings were seen for the separately expressed protein fragment hDpl(52–152), which corresponds to the globular domain of hDpl, indicating that the observed doublings are a property of the globular domain of hDpl and cannot be attributed to different interactions with the tail.

Considering the aforementioned observations on the "doubling" of part of the NMR spectrum, we used the following procedure for obtaining a single set of chemical shift assignments for the structure calculation. (i) For doubled peaks separated by less than 0.02 ppm in the ^{1}H-dimension and/or less than 0.1 ppm in the ^{15}N or ^{13}C dimension of the heteronuclear-resolved 3D [^{1}H,^{1}H]-NOESY spectra (Supplementary Material, Figure S1), only one chemical shift in each dimension was assigned and the sum of the peak intensities was used. (ii) For doubled resonances separated by more than these limits, the more intense peak was arbitrarily added to the input, with an intensity corresponding to the sum of the intensities of the two peaks. (iii) Once approximately 80% of the peaks in the 3D [^{1}H,^{1}H]-NOESY spectra had thus been picked with the program XEASY,[22] a first round of NOE assignment and structure calculation was performed using the programs CANDID[23] and DYANA.[24] The correct global fold was thus obtained, but there was locally erroneous packing of the peptide segments near the helix α2 and the β1-strand. (iv) The CANDID NOE assignments were assessed against the manually prepared peak lists, some identifications of doubled peaks were exchanged, and additional assignments were obtained by reference to the intermediate structure in the next round of structure calculation with CANDID/DYANA. This was repeated through several additional rounds of calculations until a self-consistent combination of peak picking, NOE assignment and 3D protein structure was obtained.

The NMR structure of hDpl(24–152)

The programs CANDID and DYANA yielded assignments for 5117 NOE cross-peaks, from which 2064 NOE upper distance limits for the major conformation of hDpl(24–152) (see above) were derived (Table 1). The structure of hDpl(24–152) (Figure 1) contains a well defined globular domain and a flexibly disordered tail of residues 24–50, as is seen readily from the relative intensities of the ^{15}N{^{1}H}-NOEs (Supplementary Material, Figure S2a). The globular domain of hDpl includes four α-helices comprising residues 72–80 (α1),

Table 1. Input for the structure calculation of hDpl(24–152) and characterization of the energy-minimized NMR structure of the globular domain 52–152

	Value[a]
NOE upper distance limits for hDpL(24–152)[b]	2084
NOE upper distance limits for hDpL(52–152)[c]	1739
Dihedral angle constraints for hDpl(52–152)[c]	84
Residual target function (Å^2)[b]	132 ± 0.41
Residual distance constraint violations[b]	
Number ≥ 0.1 (Å)	30 ± 9
Maximum (Å)	0.14 ± 0.03
Residual dihedral angle constraint violation[b]	
Number ≥ 2.0 deg.	0 ± 0
Maximum (deg.)	1.40 ± 0.66
AMBER energies (kcal/mol)[b]	
Total	−4841 ± 83
Van der Waals	−292 ± 14
Electrostatic	−5669 ± 79
rmsd from ideal geometry[b]	
Bond lengths (Å)	0.0078 ± 0.0002
Bond angles (deg.)	2.02 ± 0.04
rmsd to the mean coordinates (Å)[b,d]	
bb (52–150)	0.90 ± 0.18
bb (52–90,101–121,126–141)	0.55 ± 0.08
ha (52–150)	1.35 ± 0.19
ha (52–90,101–121,126–141)	0.98 ± 0.10

[a] Except for the top three entries, the average for the 20 energy-minimized conformers with the lowest residual DYANA target function values and the standard deviation among them are given.
[b] The structure calculation and the energy-minimization were performed with the polypeptide fragment of residues 24–152, and these parameters relate to this calculation.
[c] Here, the numbers of constraints within the globular domain of hDpl(52–152) are given.
[d] bb indicates the backbone atoms, N, C^α, and C′; ha stands for "all heavy atoms." The numbers in parantheses indicate the residues for which the rsmd was calculated, which are in all instances part of the globular domain.

101–115 (α2[a]), 117–121 (α2[b]) and 127–141 (α3). There is also a short two-stranded anti-parallel β-sheet comprising residues 58–60 (β1) and 88–90 (β2), which was identified on the basis of strong H^α(F59)–H^α(I89), H^α(G88)–H^N(I89) and H^α(A58)–H^N(F59) NOEs,[21] with the inter-strand hydrogen bonds H^N(H90)–O′(A58) and H^N(I60)–O′(G88) implicated by the atom coordinates of the energy-refined structure. The superposition of the backbone N, C^α and C′ atoms of the energy-minimized 20 conformers with the lowest DYANA target function values (Figure 2(a)) shows great precision of the structure determination for the peptide segments with regular secondary structure, which enclose a well-defined hydrophobic core (Figure 2(b) and (c)). The helix α1 is involved in tight hydrophobic contacts of F70, Y77, Y78 and F85 with the aliphatic side-chains of L127, V131 and L135 in helix α3. The side-chain of F59 in the β1-strand has long-range contacts to P86 and I89 of the β2-strand, T112 in the helix α2, F120 in the helix α2[b], and L132 and V136 in the helix α3. The hydrophilic side-chain of T112 is part of the core of hDpl, with its hydroxyl proton forming a long-range hydrogen bond with O′ of G57, and only 0.2% of the surface area of T112 being solvent-accessible. In addition to the insertion of the side-chain of F120 into the hydrophobic core of hDpl, the positioning of α2[b] relative to the helices α2[a] and α3 is supported by the implication from the atomic coordinates in the energy-refined structure of a hydrogen bond H^ε(Q117)–O′(113Q).

The C-terminal peptide segment of residues 143–152 adopts a non-regular secondary structure, which is folded back against the loop connecting β2 with α2. This region of the structure is globally defined only poorly (Figure 2(a)), and the local rmsd values are increased significantly for these polypeptide segments (Supplementary Material, Figure S2c). This disordered region of hDpl includes the disulfide bond between C94 and C145. When compared to the remainder of the globular domain, there is a low propensity of medium-range and long-range NOE constraints in this region (Supplementary Material, Figure S2b), but the observation of positive $^{15}N\{^1H\}$-NOEs for all amide moieties that were assigned in this region indicates that there is static disorder rather than increased intramolecular mobility.

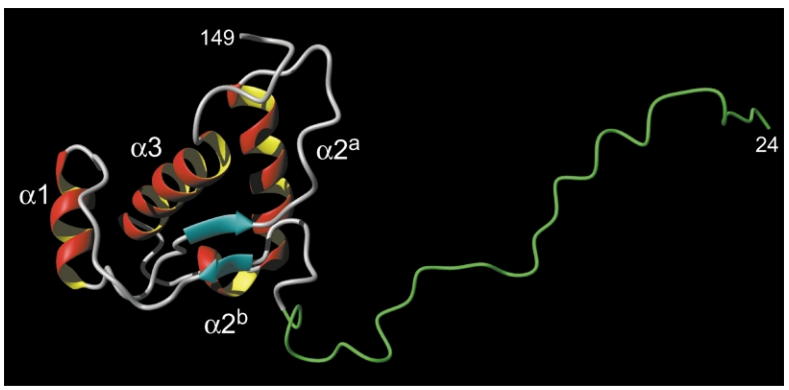

Figure 1. A cartoon of the three-dimensional structure of the recombinant full-length human doppel protein, hDpl(24–152). Helical secondary structures are red and yellow, β-strands are cyan, the segments with non-regular secondary structure within the C-terminal globular domain are gray, and the flexibly disordered N-terminal "tail" of residues 24–51 is represented by a green line, with a length approximately to scale with the globular domain. The two chain ends and the α-helices are identified.

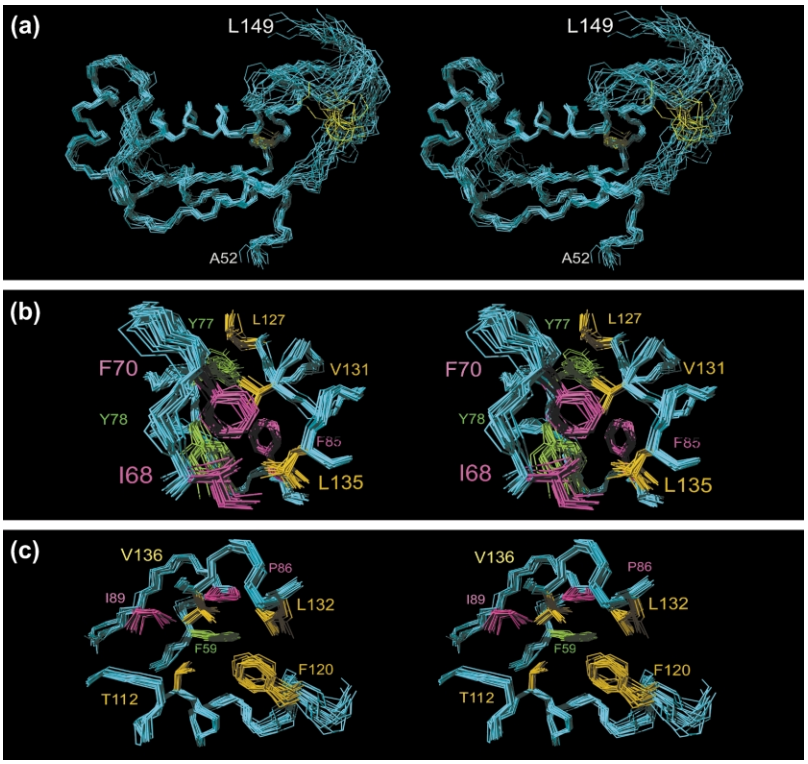

Figure 2. Stereo drawings of regions of the globular domain of hDpl(24–152). (a) Bundle of 20 energy-minimized conformers of the polypeptide segment 52–149 of hDpl(24–152) obtained by superposition of the N, C^{α} and C' atoms of the residues 52–90, 101–121 and 126–141 for best fit. The backbone (cyan), and the two disulfide bridges (yellow) are displayed. (b) Packing of amino acid side-chains in the core of hDpl(24–152). The residues shown originate from loops preceding and following helix α1 (magenta), helix α1 (green), and helix α3 (orange). (c) Environment of the residue F59, shown in green, formed by residues in the helices α2 and α3 (orange), and in the first β-strand and the preceding loop (magenta). In (b) and (c) the residue types and the sequence numbers are indicated next to the corresponding residues, using the same color code as for the amino acid side-chains, and the backbone is shown in cyan.

Discussion

Comparison of the human and murine doppel proteins

The structure alignment of hDpl(24–152) with mDpl(26–157) in Figure 3 was based primarily on a superposition of the 3D structures (Figure 2(a))[19] for best fit of the backbone heavy atoms of the peptide segments 56–91, 100–117 and 127–142 in hDpl with the corresponding segments of mDpl (Figure 3), which yielded an rmsd value of 1.8 Å (Figure 4). This alignment differs in some details from the previously used, entirely sequence-based alignment.[11,19] A corresponding best fit calculation with structure superposition only for the hDpl peptide segments 56–91 and 127–142, which excludes the helices $\alpha 2^{a}$ and $\alpha 2^{b}$, yielded an rmsd value of 0.7 Å. Although the structures of human and murine Dpl are globally very similar (Figure 4), there are thus indications of significant local structure differences.

There are 22 amino acid substitutions between the globular domains of human and murine Dpl

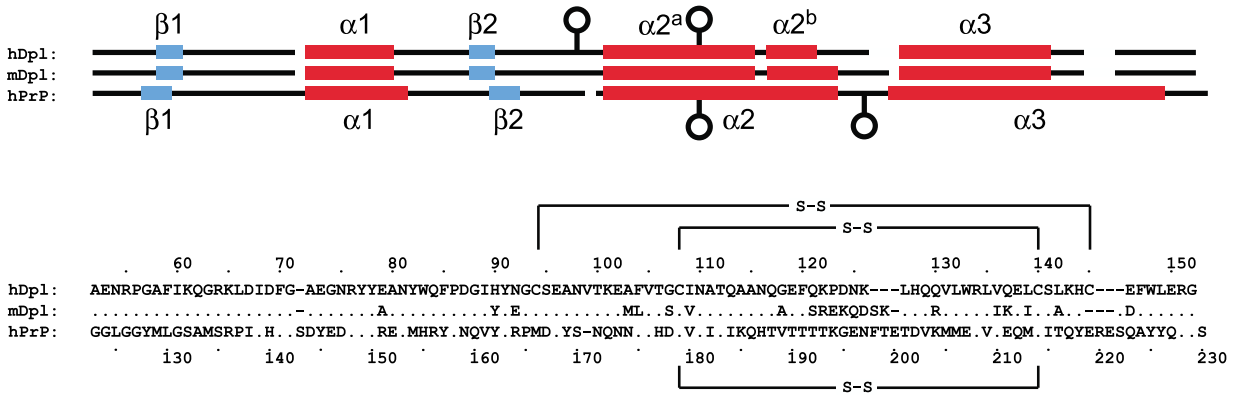

Figure 3. Tentative alignment (see the text) of the regular secondary structures (top) and the amino acid sequences (bottom) of human Dpl, murine Dpl and human PrP for polypeptide segments corresponding to the hDpl globular domain. The complete sequence is given for hDpl. For the other proteins, sequence positions that are conserved relative to hDpl are indicated by dots. Apparent deletions are indicated by a hyphen. Cirlces in the drawing of the secondary structure indicate N-glycosylation sites, which are conserved between hDpl and mDpl. The disulfide bonds are indicated. The pairwise amino acid identities between hDpl/mDpl, hDpl/hPrP and mDpl/hPrP are 79%, 20% and 20%, respectively.

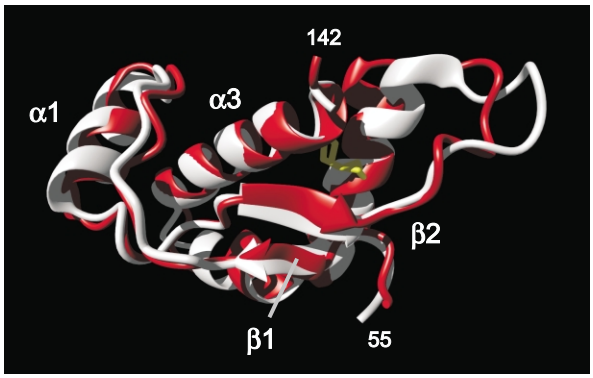

Figure 4. Superposition of the well-structured fragment 55–142 of hDpl(24–152) (red) and the corresponding fragment of mDpl(26–155) (white) (Figure 3), for best fit of the backbone heavy atoms of the residues 56–91, 100–117 and 127–142 in hDpl with the corresponding residues in mDpl. For each protein the conformer with the smallest rmsd to the mean of the final 20 conformers is shown (mDpl: PDB entry code 1I17). Regular secondary structure elements are indicated next to the ribbons, and the chain ends are indicated by sequence numbers. The disulfide bridge homologous to that in the prion protein is shown in yellow (the second disulfide bond is outside of the well-structured region of the protein, see Figure 2(a)).

(Figure 3), corresponding to an amino acid sequence conservation of 79%. Nine of these substitutions are on the surface of the molecule, i.e. E79A, H90Y, N92E, G107S, G118A, Q130R, Q137K, L142A and E146D. None of these side-chains appears to be involved in long-range or medium-range interactions with other segments of the polypeptide chain, so that they are unlikely to contribute significantly to structural differences between the two proteins. In the helix $\alpha2^b$ and the loop connecting $\alpha2^b$ with α3 there are nine differences between the two sequences, i.e. the amino acid exchanges G118A, Q121S, K122R, P123E, D124K, N125Q and K126D, and a dipeptide insertion following position 126 in the hDpl amino acid sequence. As a consequence, the helix $\alpha2^b$ is extended by about one turn in mDpl. In hDpl, P123 is at the end of $\alpha2^b$, where it initiates the loop connecting $\alpha2^b$ and α3. Interestingly, although four variations of the amino acid sequence near helix $\alpha2^b$ involve charged residues, the net local charge of the segment 117–127 (hDpl numeration) is invariant between the two proteins. The remaining four amino acid exchanges, F104L, I109V, V136I and L139A, are in the protein core. The substitutions I109V and V136I strictly preserve the hydrophobic core packing, as they form complementary surfaces on the helices α2 and α3, whereas the substitutions F104L and L139A could relate to the structural differences between the two proteins at the start of helix $\alpha2^a$ (Figure 4).

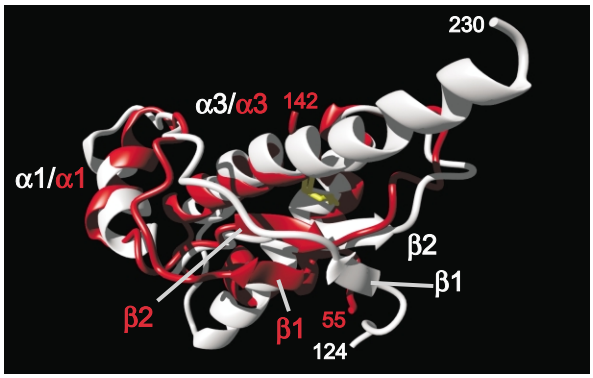

Figure 5. Superposition of the fragment 55–142 of hDpl (red) and the fragment 124–230 of hPrP (white) for best fit of the backbone heavy atoms of the hPrP polypeptide segments 139–142, 144–170, 171–186 and 200–215 with the hDpl segments 68–71, 72–98, 100–115 and 126–141. For each protein the conformer with the smallest rmsd to the mean of the final 20 conformers is shown (hPrP: PDB entry code 1QM0). Regular secondary structure elements are indicated next to the ribbons, and the chain ends are indicated by sequence numbers. The homologous disulfide bridge is shown in yellow.

It has been suggested, on the basis of the high degree of amino acid conservation, that all mammalian Dpl proteins should have a very similar 3D structure.[11,19] From the present data, we can add the prediction that other mammalian Dpl proteins will probably be closer to the 3D structure of mDpl than to the 3D structure of hDpl, since the amino acid sequences of mDpl and other non-human Dpl proteins are virtually identical in the region of helix $\alpha2^b$.

Comparison of the human doppel protein and the human prion protein

The amino acid conservation between the globular domains of hDpl and hPrP is 20%, i.e. there are 20 conserved residues. The tentative sequence alignment in Figure 3, which differs in details from a previously used, entirely sequence-based alignment,[11,19] is based on the observation that conserved residues are found in similar relative positions of the helical secondary structures and include the disulfide bond between α2 and α3, and is largely the result of a 3D structure superposition for best fit. Thereby, the similar relative orientations of the three α-helices and the β2-strand were used as the initial references, and in the final analysis the structurally corresponding peptide segments 68–71, 72–98, 100–115 and 126–141 in Dpl, and 139–142, 144–170, 171–186 and 200–215 in hPrP were identified. A global superposition obtained as the best fit for the backbone heavy atoms of these peptide segments (Figure 5) yielded an rmsd value of 2.5 Å. Figure 5 shows quite close similarity between the two proteins in the region of α1 and the sequentially adjoining residues, and in the N-terminal halves of the helices α2 and α3. The preservation of this scaffold, which shows above-average local sequence identity, indicates a common PrP-type fold for hPrP and hDpl, although there are extensive differences in the C-terminal halves of the

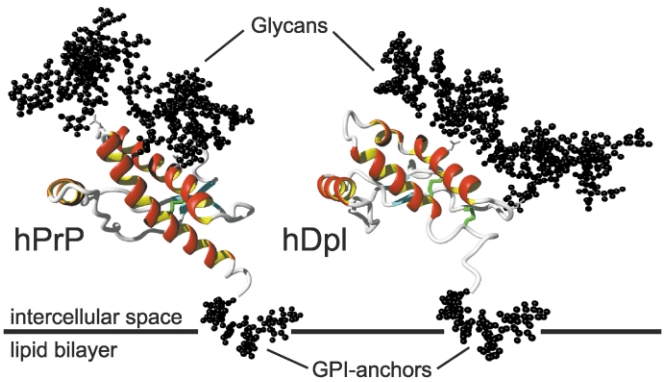

Figure 6. A cartoon of glycosylated, GPI-anchored hPrP and hDpl. The α-helices are red and yellow, the β-strands are cyan, and the segments with non-regular secondary structure within the C-terminal domain are gray. The GPI anchor and the glycan moieties are black, and the disulfide bridges are green. For simplicity, the flexibly disordered N-terminal tails are not drawn. The cell-surface membrane is indicated by a horizontal black line.

helices α2 and α3, and in the locations of the β-sheet relative to the core of the protein.

Most of the conserved residues can readily be attributed structural roles in the molecular architectures of hDpl and hPrP. The residues I68 (I139 in hPrP), F70 (F141), Y77 (Y149) and Y78 (Y150) constitute the interior hydrophobic surface of helix α1, which is packed against helix α3. The residue G71 (G142), which precedes the helix α1 in both molecules, has a Gly-accessible but otherwise disfavored backbone Φ-angle of about 100°, and appears to be required in the tight turn that initiates helix α1. Residue N81 (N153) forms a C-capping box, and thus stabilizes helix α1 at its C-terminal end, and P86 (P158) causes a bend in the polypeptide chain that initiates the β2-strand. The conserved residue Y91 (Y163) is largely solvent-inacessible in both proteins, indicating that it is involved in maintaining proper packing of part of the molecular core. The residues F104 (F175), V105 (V176) and C108 (C179) initiate the second α-helix, mediate a conserved disulfide bond between α2 and α3, and contact with V136 (V210) in α3. The side-chain of T112 (T183) in α2 forms a long-range hydrogen bond to the β-sheet, whereby in hDpl this interaction involves the hydroxyl proton of T112 and $O^{\prime}$ of G57 on strand β1, while in hPrP it involves the hydroxyl oxygen atom of T183 and H^{N} of Y162 on strand β2. This switch in donor/acceptor function is reflected in largely different χ^{1} angles of this Thr residue, which is virtually solvent-inaccessible in both hDpl and hPrP. In hPrP, this core hydrogen bond has been shown to contribute significantly to the thermodynamic stability of PrP,[25,26] and it probably has a similar role in Dpl. A conserved glycosylation site between hDpl and hPrP (Figures 3 and 6) is found near C108 (C179) in the conserved disulfide-bond, at position N110 (N181), which is part of an NXT N-glycosylation signal-sequence formed by residues 110–112.

In the superposition of Figure 5, the β1-strand and the sequentially adjacent residues show a displacement of corresponding C^{α} atom coordinates of approximately 8 Å. The β-sheet is shifted by two residues towards helix α1, resulting in a significant rearrangement of the loop connecting β1 and α1 with respect to helix α3, while leaving the position of the loop connecting α1 and β2 virtually unaffected. There is also different packing of the side-chains in this molecular region, where residue F59 of the β1-strand in hDpl is part of the hydrophobic core (Figure 2(c)). The C-terminal polypeptide segments of hPrP and hDpl have a low level of sequence conservation and different 3D structures. While in hPrP the helix α3 proceeds almost to the C terminus, the helix α3 of Dpl terminates after C140 in the common disulfide bond. The peptide segment between the end of α3 and the chain end of hDpl has a non-regular secondary structure and is folded against the loop connecting β2 with α2 (Figure 1), giving the hDpl molecule an overall more contracted appearance when compared with hPrP. It is noteworthy that the additional disulfide bond C94–C145 in hDpl would be sterically incompatible with a regular α-helix structure beyond about residue 142.

Implications for function in health and disease

The cellular functions of hDpl and hPrP in the healthy organism are unknown, which limits the extent to which structure–function correlations can be evaluated. Since there is evidence for an evolutionary relationship between the two proteins (Figure 3), the following empirical observations may be of interest for future work.

There are four conserved residues, i.e. R76 (R148 in hPrP), E96 (E168), N110 (N183) and R134 (R208), with no apparent structural roles in either of the two proteins. They do not take part in any hydrogen bonds in the energy-minimized structures of either hDpl or hPrP.[17]

The surface of hDpl contains a hydrophobic cleft, which is surrounded by charged residues (Figure 7). This cleft is sufficiently spacious to accommodate large hydrophobic side-chains, such as an indole ring of tryptophan. The residues forming this surface are highly conserved in doppel proteins from different species. Considering that this feature is conserved evolutionarily, it might represent a binding site for an unidentified, functionally important factor. Since no similar surface area is

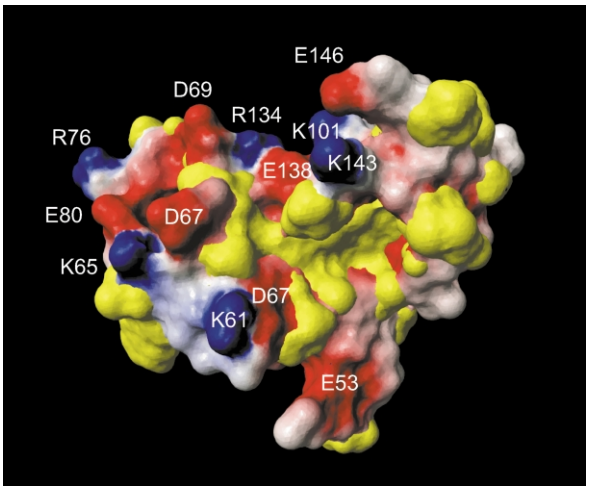

Figure 7. Surface view of hDpl in the orientation of the molecule shown in the ribbon diagram of Figure 5. Negative charges are shown in red, positive charges in blue, and residues belonging to the hydrophobic core and other hydrophobic side-chains are shown in yellow. The charged residues are identified by the one-letter symbols and the sequence numbers.

observed in PrP, it might support a function that is unique to doppel proteins.

In the pathology of TSEs the refolding of PrP^{C} to PrP^{Sc} appears to be a central event. Substitutions of the residues D178, V180, T183, R208, V210, E211 and Q212 in hPrP, which correspond to G107, I109, T112, R134, V136, Q137 and E138 in hDpl have been identified as genetic mutations that relate to increased probability to develop familial forms of TSE. Three of these residues are conserved in the sequences of hDpl and hPrP, i.e. T112 (T183), R134 (R208) and V136 (V210), and are found in corresponding locations in the 3D structures.

The second disulfide bond of hDpl, with the residues C94 and C145, connects the loop between β2 and α2 with the C-terminal peptide segment. The sequential spacing of C140 and C145 would not be compatible with an α-helical secondary structure after residue 142 of hDpl. On the one hand, the resulting non-regular secondary structure in this region contributes to the aforementioned hydrophobic surface cleft. On the other hand, combined with the second N-glycosylation site of hDpl in the loop connecting β2 and α2 (Figures 3 and 6) it causes a rather dramatic structural difference when compared to the corresponding region of hPrP. Since this loop in hPrP has been suggested to be part of an epitope for binding of an unidentified "protein X" that would influence inter-species transmissibility of TSEs,[27,28] this structural difference might be a basis for crucial functional differences between the two proteins. In particular, it seems unlikely that competition of Dpl and PrP for a common interaction site[29,30] would involve this protein X.

Materials and Methods

The protein-coding region of the Dpl-Gene, *Prnd*, was amplified from human genomic DNA (source: T. Lührs) using the primers CGGATCCGCTGAGAACCGCCCGG-GAGCCTT and CGAATTCTTAGCCCCTCTCCAACCA-AAACTC. PCR products were restricted using *Bam* HI and *Eco* RI, and subsequently cloned into the *Escherichia coli* expression vector pRSETA, which also codes for a 17 residue N-terminal histidine-tail and an engineered thrombin cleavage site.[31] The expression product, hDpl(24–152), was verified by DNA-sequencing, N-terminal Edman sequencing and MALDI-TOF mass spectroscopy.

Single clones of freshly transformed *E. coli* cells were inoculated either into Luria Bertani (LB) medium or into M9-minimal medium containing [^{15}N]ammonium chloride and [$^{13}C_6$]glucose as the only sources of nitrogen and carbon, respectively. The cells were incubated at 37 °C and shaken vigorously until an A_{600} of 0.8–1.0 was reached. The cultures were induced by adding isopropyl-β,D-thiogalactoside (IPTG) to a final concentration of 1 mM, incubated for another four to six hours at 37 °C, and the cells were then harvested by centrifugation.

Cells harvested from 4 l of culture were resuspended in buffer G (6 M GdHCl, 100 mM NaPi, 10 mM Tris (pH 8.0), 10 mM reduced glutathione (GSH), 10 mM imidazole). The suspension was sonicated for six minutes, using pulses of 0.5 second that were interspaced by 1.5 seconds. Undissolved material was sedimented at 20,000*g* for 30 minutes, and the supernatant was incubated for one hour with Ni-NTA agarose. The agarose was washed several times with buffer G. The protein was refolded while bound to the agarose, using a linear gradient of 0–100% (v/v) of buffer B (100 mM NaPi, 10 mM Tris (pH 8.0), 10 mM imidazole). The refolded hDpl(24–152) was then eluted with buffer B containing 500 mM imidazole. To maximize the protein yield, up to five refolding cycles were performed. The N-terminal histidine tail was removed by digestion with 0.05 unit of thrombin/ml in 5 mM Tris at pH 8.2. The reaction was stopped after incubation at room temperature for two hours by the addition of phenylmethylsulfonyl fluoride (PMSF) to a final concentration of 1 mM. The cleavage products were separated by cation-exchange chromatography using CM52-cellulose, from which the protein was eluted with a linear gradient of 0–500 mM sodium chloride in 10 mM Tris at pH 8.2. Protein-containing fractions were combined, and dialyzed exhaustively. Finally, purified hDpl(24–152) was concentrated to 0.5–1.0 mM, and either lyophilized for storage or used as a fresh solution.

NMR spectroscopy and structure determination

For the structure determination, we used a 1.3 mM solution of uniformly ^{13}C,^{15}N-labeled protein in a mixed solvent of 95% (v/v) H_2O, 5% (v/v) ^2H_2O and a 1.9 mM solution of unlabeled protein in 100% ^2H_2O. The NMR samples contained 10 mM sodium acetate at pH 4.5. Details on the NMR experiments used and the structure determination are given as Supplementary Material. Methods described in references 32–39 have been used.

Data Bank entries

The 20 energy-minimized conformers of the globular domain of hDpl with the lowest DYANA target function

values in the final structure calculation of hDpl(24–152) were deposited in the Protein Data Bank (PDB-entry code 1LG4). The chemical shift list for hDpl(24–152) has been deposited in the BioMagResBank (entry 5145).

Acknowledgements

We thank Dr Markus Glatzel for taking blood from T. Lührs for the preparation of human genomic DNA. This work was supported by the Schweizerischer Nationalfonds (project 31-49047.96) and the National Center of Competence in Research (NCCR) Structural Biology.

References

1. Oesch, B., Westaway, D., Walchli, M., McKinley, M. P., Kent, S. B., Aebersold, R. *et al.* (1985). A cellular gene encodes scrapie PrP 27-30 protein. *Cell*, **40**, 735–746.
2. Basler, K., Oesch, B., Scott, M., Westaway, D., Walchli, M., Groth, D. F. *et al.* (1986). Scrapie and cellular PrP isoforms are encoded by the same chromosomal gene. *Cell*, **46**, 417–428.
3. Prusiner, S. B. (1998). Prions. *Proc. Natl Acad. Sci. USA*, **95**, 13363–13383.
4. Büeler, H., Aguzzi, A., Sailer, A., Greiner, R. A., Autenried, P., Aguet, M. *et al.* (1993). Mice devoid of PrP are resistant to scrapie. *Cell*, **73**, 1339–1347.
5. Fischer, M., Rülicke, T., Raeber, A., Sailer, A., Moser, M., Oesch, B. *et al.* (1996). Prion protein (PrP) with amino-proximal deletions restoring susceptibility of PrP knockout mice to scrapie. *EMBO J.* **15**, 1255–1264.
6. Flechsig, E., Shmerling, D., Hegyi, I., Raeber, A. J., Fischer, M., Cozzio, A. *et al.* (2000). Prion protein devoid of the octapeptide repeat region restores susceptibility to scrapie in PrP knockout mice. *Neuron*, **27**, 399–408.
7. Manson, J. C., Clarke, A. R., Hooper, M. L., Aitchison, L., McConnell, I. & Hope, J. (1994). 129/Ola mice carrying a null mutation in PrP that abolishes mRNA production are developmentally normal. *Mol. Neurobiol.* **8**, 121–127.
8. Sakaguchi, S., Katamine, S., Shigematsu, K., Nakatani, A., Moriuchi, R., Nishida, N. *et al.* (1995). Accumulation of proteinase K-resistant prion protein (PrP) is restricted by the expression level of normal PrP in mice inoculated with a mouse-adapted strain of the Creutzfeldt-Jakob disease agent. *J. Virol.* **69**, 7586–7592.
9. Sakaguchi, S., Katamine, S., Nishida, N., Moriuchi, R., Shigematsu, K., Sugimoto, T. *et al.* (1996). Loss of cerebellar Purkinje cells in aged mice homozygous for a disrupted PrP gene. *Nature*, **380**, 528–531.
10. Moore, R. (1997). Gene targeting studies at the mouse prion protein locus, PhD thesis, University of Edinburgh, Edinburgh, Scotland
11. Moore, R. C., Lee, I. Y., Silverman, G. L., Harrison, P. M., Strome, R., Heinrich, C. *et al.* (1999). Ataxia in prion protein (PrP)-deficient mice is associated with upregulation of the novel PrP-like protein doppel. *J. Mol. Biol.* **292**, 797–817.
12. Behrens, A., Brandner, S., Genoud, N. & Aguzzi, A. (2001). Normal neurogenesis and scrapie pathogenesis in neural grafts lacking the prion protein homologue Doppel. *EMBO Rep.* **2**, 347–352.
13. Riek, R., Hornemann, S., Wider, G., Billeter, M., Glockshuber, R. & Wüthrich, K. (1996). NMR structure of the mouse prion protein domain PrP(121-321). *Nature*, **382**, 180–182.
14. Riek, R., Hornemann, S., Wider, G., Glockshuber, R. & Wüthrich, K. (1997). NMR characterization of the full-length recombinant murine prion protein, mPrP(23-231). *FEBS Letters*, **413**, 282–288.
15. James, T. L., Liu, H., Ulyanov, N. B., Farr-Jones, S., Zhang, H., Donne, D. G. *et al.* (1997). Solution structure of a 142-residue recombinant prion protein corresponding to the infectious fragment of the scrapie isoform. *Proc. Natl Acad. Sci. USA*, **94**, 10086–10091.
16. Liu, H., Farr-Jones, S., Ulyanov, N. B., Llinas, M., Marqusee, S., Groth, D. *et al.* (1999). Solution structure of Syrian hamster prion protein rPrP(90-231). *Biochemistry*, **38**, 5362–5377.
17. Zahn, R., Liu, A., Lührs, T., Riek, R., von Schroetter, C., Lopez Garcia, F. *et al.* (2000). NMR solution structure of the human prion protein. *Proc. Natl Acad. Sci. USA*, **97**, 145–150.
18. Lopez Garcia, F., Zahn, R., Riek, R. & Wüthrich, K. (2000). NMR structure of the bovine prion protein. *Proc. Natl Acad. Sci. USA*, **97**, 8334–8339.
19. Mo, H., Moore, R. C., Cohen, F. E., Westaway, D., Prusiner, S. B., Wright, P. E. *et al.* (2001). Two different neurodegenerative diseases caused by proteins with similar structures. *Proc. Natl Acad. Sci. USA*, **98**, 2352–2357.
20. Grathwohl, C. & Wüthrich, K. (1976). The X-Pro peptide bond as an NMR probe for conformational studies of flexible linear peptides. *Biopolymers*, **15**, 2025–2041.
21. Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, Wiley, New York.
22. Bartels, C., Xia, T. H., Billeter, M., Güntert, P. & Wüthrich, K. (1995). The program XEASY for computer-supported NMR spectral-analysis of biological macromolecules. *J. Biomol. NMR*, **6**, 1–10.
23. Herrmann, T., Güntert, P. & Wüthrich, K. (2002). Protein NMR structure determination with automated NOE assignment using the new softwear CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319**, 209–227.
24. Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298.
25. Liemann, S. & Glockshuber, R. (1999). Influence of amino acid substitutions related to inherited human prion diseases on the thermodynamic stability of the cellular prion protein. *Biochemistry*, **38**, 3258–3267.
26. Riek, R., Wider, G., Billeter, M., Hornemann, S., Glockshuber, R. & Wüthrich, K. (1998). Prion protein NMR structure and familial human spongiform encephalopathies. *Proc. Natl Acad. Sci. USA*, **95**, 11667–11672.
27. Telling, G. C., Scott, M., Mastrianni, J., Gabizon, R., Torchia, M., Cohen, F. E. *et al.* (1995). Prion propagation in mice expressing human and chimeric PrP transgenes implicates the interaction of cellular PrP with another protein. *Cell*, **83**, 79–90.
28. Scott, M., Groth, D., Foster, D., Torchia, M., Yang, S. L., DeArmond, S. J. *et al.* (1993). Propagation of

prions with artificial properties in transgenic mice expressing chimeric PrP genes. *Cell*, **73**, 979–988.
29. Shmerling, D., Hegyi, I., Fischer, M., Blattler, T., Brandner, S., Gotz, J. *et al.* (1998). Expression of amino-terminally truncated PrP in the mouse leading to ataxia and specific cerebellar lesions. *Cell*, **93**, 203–214.
30. Rossi, D., Cozzio, A., Flechsig, E., Klein, M. A., Rülicke, T., Aguzzi, A. *et al.* (2001). Onset of ataxia and Purkinje cell loss in PrP null mice inversely correlated with Dpl level in brain. *EMBO J.* **20**, 694–702.
31. Zahn, R., von Schroetter, C. & Wüthrich, K. (1997). Human prion proteins expressed in *Escherichia coli* and purified by high-affinity column refolding. *FEBS Letters*, **417**, 400–404.
32. Bax, A. & Grzesiek, S. (1993). Methodological advances in protein NMR. *Accts Chem. Res.* **26**, 131–138.
33. Ikura, M., Bax, A., Clore, G. M. & Gronenborn, A. M. (1990). Detection of nuclear Overhauser effects between degenerate amide proton resonances by heteronuclear 3-dimensional nuclear-magnetic-resonance spectroscopy. *J. Am. Chem. Soc.* **112**, 9020–9022.
34. Boelens, R., Burgering, M., Fogh, R. H. & Kaptein, R. (1994). Time-saving methods for heteronuclear multidimensional NMR of (C-13,N-15) doubly labeled proteins. *J. Biomol. NMR*, **4**, 201–213.
35. Kumar, A., Ernst, R. R. & Wüthrich, K. (1980). A two-dimensional nuclear Overhauser enhancement (2d Noe) experiment for the elucidation of complete proton–proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* **95**, 1–6.
36. Koradi, R., Billeter, M. & Güntert, P. (2000). Point-centered domain decomposition for parallel molecular dynamics calculation. *Comput. Phys. Commun.* **124**, 139–147.
37. Luginbühl, P., Güntert, P., Billeter, M. & Wüthrich, K. (1996). The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. *J. Biomol. NMR*, **8**, 136–146.
38. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M. *et al.* (1995). A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J. Am. Chem. Soc.* **117**, 5179–5197.
39. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55.

Edited by M. F. Summers

(Received 14 June 2002; received in revised form 25 November 2002; accepted 25 November 2002)

Supplementary Material comprising details on the NMR experiments used and two Figures is available on Science Direct