Leptin is a four-helix bundle: secondary structure by NMR

Allen D. Kline*, Gerald W. Becker, Lisa M. Churgay, Bryan E. Landen, Debra K. Martin, William L. Muth, Radhakrishnan Rathnachalam, John M. Richardson, Brigitte Schoner, Maverick Ulmer, John E. Hale

Lilly Research Laboratories, Lilly Corporate Center, Indianapolis, IN 48285-0403, USA

Received 19 February 1997

Abstract Leptin is a signaling protein that in its mutant forms has been associated with obesity and Type II diabetes. The lack of sequence similarity has precluded analogies based on structural resemblance to known systems. Backbone NMR signals for mouse leptin ($^{13}C/^{15}N$ -labeled) have been assigned and its secondary structure reveals it to be a four-helix bundle cytokine. Helix lengths and disulfide pattern are in agreement with leptin as a member of the short-helix cytokine family. A three-dimensional model was built verifying the mechanical consistency of the identified elements with a short-helix cytokine core.

Key words: Nuclear magnetic resonance; Leptin; Obesity; Secondary structure; Cytokine-fold

1. Introduction

The obese gene encodes a 167-amino-acid polypeptide [1] that appears to function in the maintenance of body weight. The product of this gene, leptin, is a secreted 146-residue protein containing a single disulfide bond. Injection of murine or human leptin into ob/ob mice leads to weight loss and correction of the diabetic state of these animals [2–4]. Leptin appears to act on receptors in the arcuate nucleus to suppress neuropeptide Y release [5]. The leptin receptor and various splice variants were cloned [6–8] and found to have sequence similarity to the class I cytokine receptor family [6]. Interestingly, the expression of the leptin receptor seems to occur in many different tissues, and recent findings have implicated leptin in reproduction [9] and hematopoiesis [10].

Leptin's role as a circulating hormone and the similarity of its receptor to cytokine receptors implicate it as a cytokine; however, it has no strong sequence similarity with any other protein, making modeling of the protein structure difficult. Threading analysis of the leptin sequence against a 3-dimensional database indicated that it is compatible with a cytokine folding pattern [11]. However, the predicted helical composition of 39% is considerably lower than the 52% estimated by CD measurements [12]. While these studies implicate this fold, they do not help to resolve the structural subfamily to which leptin belongs. In this paper we will use NMR methodologies to determine the exact secondary structure for mouse leptin, giving direct experimental verification to its adoption of a cytokine fold and its similarity to the short-helix subfamily of cytokine folds.

*Corresponding author. Fax: (317) 276-9722.
E-mail: a.kline@lilly.com

Abbreviations: CD, circular dichroism; CSI, chemical shift index; IL-6, interleukin-6; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect

2. Materials and methods

E. coli expressing murine leptin was grown at 37°C overnight in either a modified M-9-defined medium enriched with ^{15}N or in medium (Isotec) enriched in ^{15}N or $^{13}C/^{15}N$ prepared according to the manufacturer's instructions. Inclusion bodies containing leptin were isolated by differential centrifugation and the protein was purified and renatured as described [5]. The isotopic enrichment of the purified protein was estimated by electrospray ionization mass spectrometry (PE-Aciex, API-III). Prior to NMR studies, the protein was dialyzed against a 10% ^2H_2O solution containing 20 mM PO_4 and 1 mM EDTA, pH 6.6, then concentrated approximately to 10 mg/ml.

NMR spectroscopy was performed on a Unity-500 (Varian Instruments) equipped with gradients. Probe temperature was regulated at 30°C. The following experiments were collected: HSQC(^{15}N), NOESY-HSQC(^{15}N), TOCSY-HSQC(^{15}N), HNCA [13], HNCACB [13], CBCA(CO)NNH [13,14], HBHA(CBCACO)NNH [13,14], HNCO [15], CBCACO(CA)HA [16], and HBHA(CBCA)NNH [14]. The experiments, starting with proton magnetization on carbon, were run with both short and long refocusing delays to distinguish the carbon's protonation state [17]. The spectra were referenced using sodium 2,2-dimethyl-2-silapentane-5-sulfonate added at the completion of data collection, referencing nitrogen and carbon dimensions indirectly [18]. Spectra were processed and analyzed using Felix-95 (Molecular Simulations). CSI was calculated with the program of Wishart and Sykes [19]. Protein modeling was completed using the Quanta program (Molecular Simulations) with coordinates obtained from the Protein Data Bank.

3. Results and discussion

For the sequential assignments of mouse leptin we utilized two labeled samples: one with ^{15}N and a second with ^{15}N and ^{13}C. Incorporation levels were determined to be 98% and 97%, respectively. During the analysis of the ^{15}N spectra, we noted more signals than expected given the amino acid composition. Also a number of NH signals were obviously doubled in a 2:1 ratio. From numerous experiments we were unable to determine the source of the heterogeneity and concluded that the isomers must be conformational (Kline, unpublished results). For this paper only the major isomer will be examined.

The assignment strategy used is based on identifying an NH intraresidual C^α, C^β, C°, H^α and H^β signals and then comparing to another NH sequential C^α, C^β, C°, H^α and H^β [20–22]. When the intraresidual and sequential signals align, then the NHs are defined as sequential. A contiguous series of NHs are examined for shift information about their amino acid type. A series is subsequently fit into the amino acid sequence.

Fig. 1 shows a summary of the scalar connections between the NHs for leptin which lead to the assignment of the major isomer. Gaps occur on this NH–NH connectivity map due to the occurrence of proline or missing NH signals (His-26, Gln-

PII S0014-5793(97)00353-0

Fig. 1. Summation of the sequential NH–NH connectivities which allowed the mouse leptin NMR assignments to be made. A black bar adjoins residues when their corresponding NH signals were connected through the indicated intervening atoms. Blanks occur either where no data were possible or where no data were observed.

28, and Gln-100). The connectivity is nearly complete for C^{α}, C^{β}, C° and H^{α} which gives confidence in the accuracy of the assignments. Fewer residues were connected through the H^{β}, due primarily to the limited number of NHs having intraresidual H^{β}. Despite this, sufficient data were available to uniquely assign all the residues. Following the completion of the bulk of the assignments, two remaining NHs, Val-1 and Thr-27, were assigned principally by elimination. A table of the assigned chemical shift values is available upon request to the authors. Finally NOE information was used to corroborate the leptin assignments.

The CSI is most accurate in determining secondary structure when a consensus analysis is performed using the four nuclei: C^{α}, C^{β}, C°, and H^{α} [19]. For leptin, these values were available for 145 of the 146 residues. The analysis (Fig. 2) reveals four long helices and one very short strand segment. There are also two relatively long random-coil loops. The positions and lengths of the helices match closely with the results from the NOE analysis. Occasionally the two methods showed slight discrepancies near the end of a helix. For instance, the shifts for Ser-67 indicate a random coil conformation, while the backbone NOEs are consistent with α-helix. These differences never extended more than a single residue.

Using the CSI and the NOEs, the secondary structure for leptin was assigned in the following manner: Helix A runs from residues 3–24, Helix B from 51–67, Helix C from 72–94, and Helix D from 122–141. Each helix contains five to six turns. Random coil occupies two long loops 25–50 and 95–121. The only β-strand occurs at 47–50, though no NOEs were identified linking it to an unidentified strand, so it is unlikely that any sheet exists in this protein. The remaining loop is a short segment that occurs from 68–71. The length of this loop and the presence of proline in the second position, is reminiscence of tight turns between β-strands [23], though this is located between helices. While the assignment of secondary structure is for the mouse protein, the same assignments are likely for the human protein as there are only five amino acid changes between the species in the helical regions [1].

Leptin's four helices and two interconnecting loops are consistent with a classification as a cytokine four-helix bundle and agrees with the earlier threading [11] and CD analysis [12]. The characteristic up-up-down-down arrangement of the helices in the cytokine fold requires long loops between helices A and B and between helices C and D [24]. The connector between helix B and C is generally shorter. The threading analysis was accurate in its identification of helical residues; however, 25 helical residues were missed because they were improperly assigned to random coil or β-strand. The CD estimate was close to the 56% measured here, but that study did not specify any of the helix positions [12].

While in general leptin's secondary structure is similar to any four-helix bundle cytokine, in specific, the length and position of its helices appears unique. The four-helix motif is able to accommodate a wide variation in helix length, so

Fig. 2. The CSI and NOE information used to determine the secondary structure for mouse leptin. The CSI histogram indicates an α-helix designation when above the line and a β-sheet when below. $d_{\alpha N}$ indicates the presence of sequential H^{α}-NH NOE. Thin line designates where $d_{\alpha N}$ was weaker than the intraresidue H^{α}-NH NOE, thick line where $d_{\alpha N}$ was stronger than the intraresidue NOE. d_{NN} indicates the presence of sequential NH–NH NOE. Thin line designates where d_{NN} was weaker than $d_{\alpha N}$, thick line where d_{NN} was stronger than $d_{\alpha N}$.

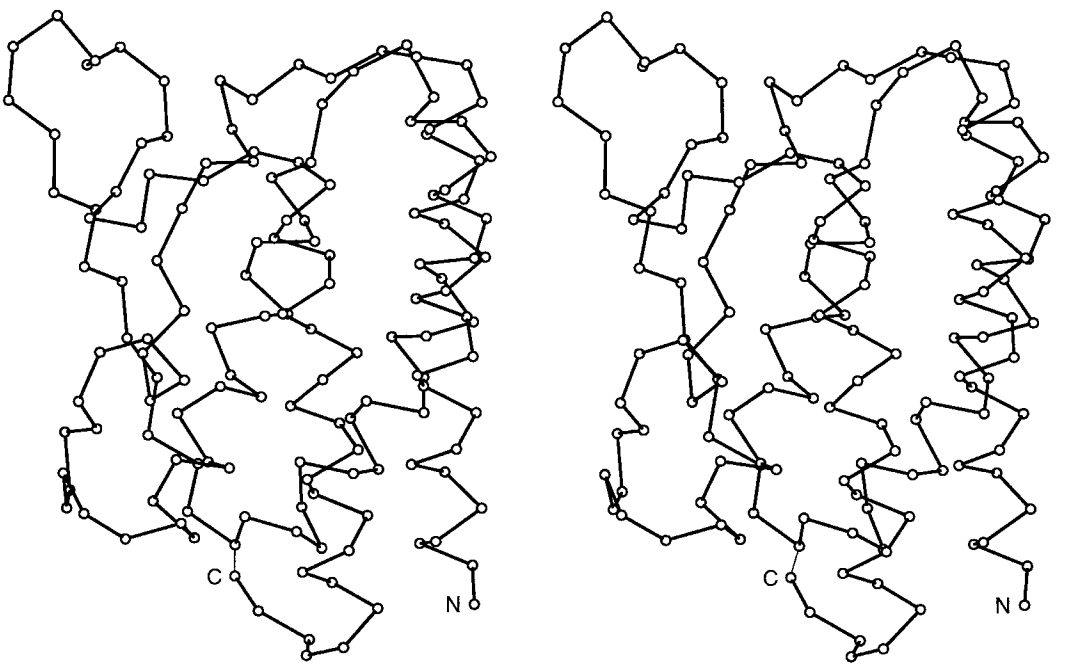

Fig. 3. A stereo C^{α} fold drawing of the mouse leptin model, based on the NMR helix assignments. C^{α} atoms are shown as open circles; termini are labeled near the bottom of the structure. The 96–146 disulfide is represented by a thin line near the C-terminus.

that an exact match would be unlikely. Leptin's helix lengths are most similar to interleukin-2 [25], interleukin-4 [26] and macrophage-colony stimulating factor [27], all members of the short-helix family [24,28]. Its single disulfide connection from the beginning of the CD loop to the C-terminus, resembles granulocyte macrophage-colony stimulating factor [29], another member of the short-helix cytokines. So from secondary structure alone, leptin is likely a member of the short-helix family. Other parameters important to classification such as loop crossover and angles between helical moments [24], will have to await the full tertiary structure.

Computer modeling was undertaken using a superposition of the helical bundles for all members of the short-helix family. The repetition of hydrophobic residues on leptin's helices matched well with the hydrophobic repeats for the known structures and was ultimately the basis for aligning the amino acid sequences for the modeling procedure. The resulting structure for leptin is shown in Fig. 3. The only significance we will draw from this model is that the four helices determined in this study fit nicely into the short-helix cytokine fold and have no difficulty making the short turn at 68–71 or the disulfide bridge at 96–146.

The length of its helices, the position of its disulfide, and the success of this modeling effort, all suggest that leptin is a member of the short-helix cytokine fold. This result is somewhat unexpected due to the receptor's sequence similarity to gp130 [30], potentially connecting it to the IL-6 family of receptors [31]. IL-6 [32] and its siblings, leukemia inhibitory factor [33] and oncostatin M [34] are all long-helix cytokines and have average helix lengths one to two turns longer than leptin. Thus while leptin and IL-6 have sequence similarities in their receptors, they do not seem to share a ligand structural classification. Recent published data also suggests that leptin is not a member of the IL-6 family because it does not use the signaling protein gp130 [35]. The leptin receptor is, however, not at all similar to the interleukin-2 family of receptors [36], thus leptin and its receptor may represent an entirely new family of Type I receptors.

Acknowledgements: The authors gratefully acknowledge Prof. Lewis E. Kay (Univ. of Toronto) for providing numerous triple resonance pulse sequences and Prof. Brian D. Sykes (Univ. of Alberta) for the automated CSI program.

References

[1] Y. Zhang, R. Proenca, M. Maffei, M. Barone, L. Leopold, J.M. Friedman, Nature 372 (1994) 425–432.
[2] M.A. Pellymounter, M.J. Cullen, M.B. Baker, R. Hecht, D. Winters, T. Boone, F. Collins, Science 269 (1995) 540–543.
[3] J.L. Halaas, K.S. Gajiwala, M. Maffei, S.L. Cohen, B.T. Chait, D. Rabinowitz, R.L. Lallone, S.K. Burley, J.M. Friedman, Science 269 (1995) 543–546.
[4] L.A. Campfield, F.J. Smith, Y. Guisez, R. Devos, P. Burn, Science 269 (1995) 546–549.
[5] T.W. Stephens, M. Basinski, P.K. Bristow, J.M. Bue-Valleskey, S.G. Burgett, L. Craft, J. Hale, J. Hoffman, H.M. Hsiung, A. Kriauciunas, W. MacKellar, P.R. Rosteck, B. Schoner, D. Smith, F.C. Tinsley, X. Zhang, M. Heiman, Nature 377 (1995) 530–532.
[6] L.A. Tartaglia, M. Dembski, X. Weng, N. Deng, J. Culpepper, R. Devos, G.J. Richards, L.A. Campfield, F.T. Clark, J. Deeds, C. Muir, S. Sanker, A. Moriarty, K.J. Moore, J.S. Smutko, G.G. Mays, E.A. Woolf, C.A. Monroe, R.I. Tepper, Cell 83 (1996) 1263–1271.
[7] C.C. Streamson, W.K. Chung, S. Wu-Peng, Y. Zhang, S. Liu, L.A. Tartaglia, R.L. Leibel, Science 271 (1996) 994–996.
[8] G. Lee, R. Proenca, J.M. Montez, K.M. Carroll, J.G. Darvishzadeh, J.I. Lee, J.M. Friedman, Nature 379 (1996) 632–635.
[9] F.F. Chehab, M.E. Lim, R. Lu, Nat Genet 12 (1996) 318–320.
[10] J.A. Cioffi, A.W. Shafer, T.J. Zupancic, J. Smith-Gbur, A. Mikhail, D. Platika, H.R. Snodgrass, Nat Med 2 (1996) 585–589.
[11] T. Madej, M.S. Boguski, S.H. Bryant, FEBS Lett 373 (1995) 13–18.
[12] F.L. Rock, S.W. Altmann, M. van Heek, R.A. Kastelein, J.F. Bazan, Horm Metab Res 28 (1996) 649–652.
[13] D.R. Muhandiram, L.E. Kay, J Magn Reson B 103 (1994) 203–216.
[14] S. Grzesiek, A. Bax, J Am Chem Soc 114 (1992) 6291.
[15] M. Wittekind, L. Mueller, J Magn Reson B 101 (1993) 201.
[16] L.E. Kay, J Am Chem Soc 115 (1993) 2055–2057.
[17] W. Feng, C. Rios, G.T. Montelione, J Biomol NMR 8 (1996) 98–104.
[18] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, B.D. Sykes, J Biomol NMR 6 (1995) 135–140.
[19] D.S. Wishart, B.D. Sykes, J Biomol NMR 4 (1994) 171–180.
[20] A. Bax, Curr Opin Struct Biol 4 (1994) 738–744.
[21] A.M. Gronenborn, G.M. Clore, Crit Rev Biochem 30 (1995) 351–385.
[22] L.E. Kay, Prog Biophys Mol Biol 63 (1995) 277–299.
[23] B.L. Sibanda, J.M. Thornton, J Mol Biol 229 (1993) 428–447.
[24] S.R. Sprang, F.J. Bazan, Curr Opin Struct Biol 3 (1993) 815–827.
[25] B.J. Brandhuber, T. Boone, W.C. Kenney, D.B. McKay, Science 238 (1987) 1707–1709.
[26] D.S. Garrett, R. Powers, C.J. March, E.A. Frieden, G.M. Clore, A.M. Gronenborn, Biochemistry 31 (1992) 4347–4353.
[27] J. Pandit, A. Bohm, J. Jancarik, R. Halenbeck, K. Koths, S.H. Kim, Science 258 (1992) 1358–1362.
[28] A.G. Murzin, S.E. Brenner, T. Hubbard, C. Chothia, J Mol Biol 247 (1995) 536–540.
[29] T. Zink, A. Ross, D. Ambrosius, R. Rudolph, T.A. Holak, FEBS Lett 314 (1992) 435–439.
[30] L.A. Tartaglia, M. Dembski, X. Weng, N. Deng, J. Culpepper, R. Devos, G.J. Richards, L.A. Campfield, F.T. Clark, J. Deeds, C. Muir, S. Sanker, A. Moriarty, K.J. Moore, J.S. Smutko, G.G. Mays, E.A. Woolf, C.A. Monroe, R.I. Tepper, Cell 83 (1995) 1263–1271.
[31] T. Kishimoto, S. Akira, T. Taga, Science 258 (1992) 593–597.
[32] G. Xu, J. Hong, T. McDonagh, M. Stahl, L.E. Kay, J. Seehra, D.A. Cumming, J Biol NMR 8 (1996) 123–135.
[33] R.C. Hoffman, F.J. Moy, V. Price, J. Richardson, D. Kaubisch, E.A. Frieden, J.D. Krakover, B.J. Castner, J. King, C.J. March, R. Powers, J Biol NMR 7 (1996) 273–282.
[34] R.C. Robinson, L.M. Grey, D. Staunton, H. Vankelecom, A.B. Vernallis, J.F. Moreau, D.I. Stuart, J.K. Heath, E.Y. Jones, Cell 77 (1994) 1101–1116.
[35] H. Baumann, K.K. Morella, D.W. White, M. Dembski, P.S. Bailon, H.K. Kim, C.F. Lai, L.A. Tartaglia, Proc Natl Acad Sci 93 (1996) 8374–8378.
[36] N. Sato, A. Miyajima, Curr Opin Cell Biol 6 (1994) 174–179.