Title

Uncovering the molecular mechanisms behind disease-associated leptin variants

Journal

Journal of Biological Chemistry, 293(33)

ISSN

0021-9258

Authors

Haglund, Ellinor
Nguyen, Lannie
Schafer, Nicholas Peter
et al.

Publication Date

2018-08-01

DOI

10.1074/jbc.ra118.003957

Peer reviewed

Uncovering the molecular mechanisms behind disease-associated leptin variants

Received for publication, May 14, 2018, and in revised form, June 14, 2018 Published, Papers in Press, June 27, 2018, DOI 10.1074/jbc.RA118.003957

Ellinor Haglund[‡1], Lannie Nguyen[§], Nicholas Peter Schafer[‡2], Heiko Lammert[‡], Patricia Ann Jennings[§3], and José Nelson Onuchic[‡¶4]

From the [‡]Center for Theoretical Biological Physics and the [¶]Departments of Physics and Astronomy, Chemistry, and Biosciences, Rice University, Houston, Texas 77005 and the [§]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093

Edited by Karen G. Fleming

The pleiotropic hormone leptin has a pivotal role in regulating energy balance by inhibiting hunger and increasing energy expenditure. Homozygous mutations found in the leptin gene are associated with extreme obesity, marked hyperphagia, and impaired immune function. Although these mutations have been characterized *in vivo*, a detailed understanding of how they affect leptin structure and function remains elusive. In the current work, we used NMR, differential scanning calorimetry, molecular dynamics simulations, and bioinformatics calculations to characterize the effects of these mutations on leptin structure and function and binding to its cognate receptor. We found that mutations identified in patients with congenital leptin deficiency not only cause leptin misfolding or aggregation, but also cause changes in the dynamics of leptin residues on the receptor-binding interface. Therefore, we infer that mutation-induced leptin deficiency may arise from several distinct mechanisms including (i) blockade of leptin receptor interface II, (ii) decreased affinity in the second step of leptin's interaction with its receptor, (iii) leptin destabilization, and (iv) unsuccessful threading through the covalent loop, leading to leptin misfolding/aggregation. We propose that this expanded framework for understanding the mechanisms underlying leptin deficiency arising from genetic mutations may be useful in designing therapeutics for leptin-associated disorders.

Obesity is one of the biggest public health challenges in the world (1), often leading to inflammation and early mortality. Furthermore, obesity increases susceptibility to other diseases, such as hyperinsulinemia, advanced bone age, hypothalamic hypothyroidism, and hypogonadotropic hypogonadism (2–4), hypothermia and cold intolerance, and hypercorticolemia and other endocrine and metabolic abnormalities (5), which lead to failure to undergo puberty and infertility (6). Although regulation of energy expenditure is considered the primary function of leptin, it also plays a role in other physiological processes, such as cancer (7), thermogenesis, and immune function (8), as evidenced by its many sites of synthesis other than adipose tissue (9) and the many cell types other than hypothalamic cells that have leptin receptors. Many of these additional functions are yet to be fully characterized.

Control of appetite and the predisposition for obesity have been linked to 41 sites on the human genome (10). One of these genes is the pleiotropic hormone leptin (11–13), which has a pivotal role in regulating energy balance through inhibition of hunger and increased energy expenditure. Interestingly, there are genetic modifications observed in patients with congenital leptin deficiency, a form of monogenic obesity characterized by its early-age onset and marked hyperphagia (14–16).

In classic leptin deficiency, no protein or very low levels of leptin are detected in serum of patients. Recently, a new type of functional leptin deficiency has been described wherein normal to high levels of bioinactive leptin are detected in the serum of patients (17). Specifically, these leptin variants are biologically inactive due to structural alteration in the receptor-binding sites, thus decreasing the affinity for its receptor, the leptin receptor (LEP-R)[5] (18, 19). These observations introduced a new form of effective leptin deficiency, wherein leptin is produced and secreted but functionally inactive because of impaired binding to LEP-R. Both classic and functional leptin deficiency can be treated by replacement with recombinant human leptin (17). Nevertheless, a more detailed mechanistic understanding of leptin's biological activity and of the different modes of failure caused by disease mutations, holds great promise for enhancing leptin's potential as a target for drug development and therapeutics.

Leptin is secreted from adipose tissue and must cross the blood–brain barrier and bind to its receptor in the hypothalamus (20) to activate the leptin signaling pathway. Additional pathways can also be activated in peripheral tissue, as the leptin

This work was supported by the Center for Theoretical Biological Physics, sponsored by National Science Foundation (NSF) Grant PHY-1427654, and NSF Grants MCB-1241332, PHY-161407, and CHE-1614101. The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This article contains Figs. S1–S3.

[1] Supported in part by Welch Foundation Grant C-1792. To whom correspondence may be addressed. E-mail: ellinorh@rice.edu.

[2] Supported by NIGMS, National Institutes of Health, Grant R01 GM44557 and the D. R. Bullard-Welch Chair at Rice University, Grant C-0016.

[3] To whom correspondence may be addressed. E-mail: pajennings@ucsd.edu.

[4] To whom correspondence may be addressed. E-mail: jonuchic@rice.edu.

[5] The abbreviations used are: LEP-R, leptin receptor; PLT, pierced lasso topology; DSC, differential scanning calorimetry; nsSNP, non-synonymous single nucleotide polymorphism; NSD, native state dynamics; RMSF, root mean square fluctuation; D3, D4, and D5, domain 3, 4, and 5, respectively; MD, molecular dynamics; HSQC, heteronuclear single quantum coherence; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol.

receptor is found on the cell surface of many organs and tissues in the body (21–25).

To activate the signaling pathway, leptin binds to the LEP-R in a sequential fashion (Fig. 1). It first makes an initial complex with the LEP-R. Once bound to leptin in a 1:1 configuration (Fig. 1*B*), the receptors can interact further and form a 2:2 quaternary complex (Fig. 1*C*) representing the core-signaling complex (26–28). Domain 3 (D3) allows for the formation of the quaternary complex that initiates signaling. Previous studies revealed that deletion of D3 completely abolished signaling (29, 30), demonstrating its integral role in formation of the quaternary complex. This is reminiscent of the role for D3 in other cytokine receptor systems, where it interacts with binding site III of the ligand already bound to another receptor chain. Once activated, the LEP-R complex signals through activation of four different signaling pathways: (i) through the Janus kinase 2 (JAK2) signaling pathway, (ii) the signal transducer and activator of transcription 3 (STAT3), (iii) the insulin receptor substrate (IRS)/phosphatidylinositol 3 kinase (PI3K), and (iv) the SH2-containing protein tyrosine phosphatase 2 (SHP2)/mitogen-activated protein kinase (MAPK) 5′-adenosine monophosphate-activated protein kinase (AMPK)/acetyl-CoA carboxylase (ACC) (31). Although the signaling complex between leptin and LEP-R is not yet fully understood, we can reconstruct some structural aspects of LEP-R activation based on (i) the structural data from Carpenter *et al.* (32), where the authors crystallize domains 4 and 5 (D4 and D5) of LEP-R together with an antibody; (ii) EM data from Mancour *et al.* (26), showing the full initial binary complex between leptin and LEP-R; and (iii) more recent work by Moharana *et al.* (27), who used SAXS experiments to investigate the activated quaternary configuration (Fig. 1*C*).

Recently, we discovered that the topology of leptin is not trivial, but instead, leptin has a "knotlike" configuration under *oxidizing* conditions (see the supporting information). Part of the polypeptide chain is threaded through the 50-residue covalent loop that is closed by the disulfide bridge formed between the two cysteines (Fig. 1*A*), resulting in a pierced lasso topology (PLT) (33). This PLT is not unique to leptin, as PLTs have been found in more than 300 proteins within the Protein Data Bank (33–39). Interestingly, we showed that the PLT is important to leptin's biological activity, as the threaded topology introduces conformational dynamics in sites far removed from the disulfide bridge. In particular, the PLT influences the dynamics of residues that are important for receptor interaction and leptin signaling (33, 35, 36, 39), thereby lowering the affinity for the LEP-R in its reduced form (33, 35). The reduced form of leptin, missing the PLT, shows decreased affinity for the LEP-R receptor (33, 35). Furthermore, the PLT has a profound effect on protein stability, as breaking the disulfide bond decreases the global stability from 3.4 to 1.8 kcal/mol (33). This is a larger perturbation of the global stability than is expected from the loss of a typical disulfide bridge that is not involved in a PLT. Given the importance of the PLT to the biological activity of leptin, the following question then arises. Is the threading event affected in the case of the disease-associated mutations identified in leptin, or can these proteins still thread into a WT-like knotlike conformation?

The recent discovery of purely functional leptin deficiency in addition to classical leptin deficiency in serum indicates that the biological function of leptin may be disrupted in multiple different ways. At the genetic level, congenital leptin deficiency is caused by homozygous mutations of either the leptin or LEP-R genes (40). In the leptin gene, 13 disease-associated mutations implicated in obesity have been found to date (see Fig. S1) (18, 19, 41–52). The affected residues are located throughout the structure of leptin, suggesting, again, that different mechanisms may be responsible for the effects of different mutations. Several *in vivo* studies of disease-associated leptin variants have been conducted, but the mechanisms that lead to defective leptin are not yet fully understood. It has been suggested that leptin's disulfide bridge is crucial for biological activity and that disruption is the cause of disease (44, 45, 48), but do conditions presented in patients with leptin deficiency simply result from the loss of a single disulfide bridge?

To elucidate the structural and functional origins of congenital leptin deficiencies that result from different homozygous mutations in leptin, we constructed and purified the seven known, full-length variants of leptin that are implicated in proteopathic diseases. These variants were probed using NMR and differential scanning calorimetry (DSC) experiments together with molecular dynamics (MD) simulations and bioinformatics techniques to evaluate the mutational effects on stability, dynamics, and solubility of the native state. As expected from previously published results (4, 41, 42, 47, 53–55), all proteins express well, suggesting that lack of protein synthesis is probably not the cause of disease. Our results, combined with previous biochemical, biological, and medicinal studies (16, 18, 19, 41–52), suggest that leptin deficiency can arise from multiple mechanisms, including (i) blockage of receptor interface through binding site II in leptin, interacting with domains 4 and 5 of the receptor to initiate a nonsignaling LEP-R complex, (ii) modifications of the affinity for domain 3 on the receptor through binding site III in leptin, perturbing the formation of the quaternary complex, leading to impaired signaling; (iii) destabilization of the native biologically active state that could occur through the loss of salt bridges or the disulfide bridge leading to an unthreaded topology and possibly degradation/clearance from serum; and (iv) unsuccessful threading through the covalent loop that might lead to protein misfolding and/or protein aggregation.

When leptin is destabilized or cannot reach the native PLT conformation, it may be either quickly degraded or may aggregate, both of which would result in impaired secretion and an absence of leptin in serum. Our data indicate that each of these mechanisms probably contributes to pathological behavior in some of the full-length disease-associated leptin variants.

Results

Known disease mutations

The gene encoding leptin is located on chromosome 7q31.3, where the translated portion of the NCBI reference sequence is NM_000230.2 (56). A sequence alignment of leptin homologs is shown as a sequence logo plot in Fig. S1 (*bottom*). The plot presents the level of conservation with the one-letter code for

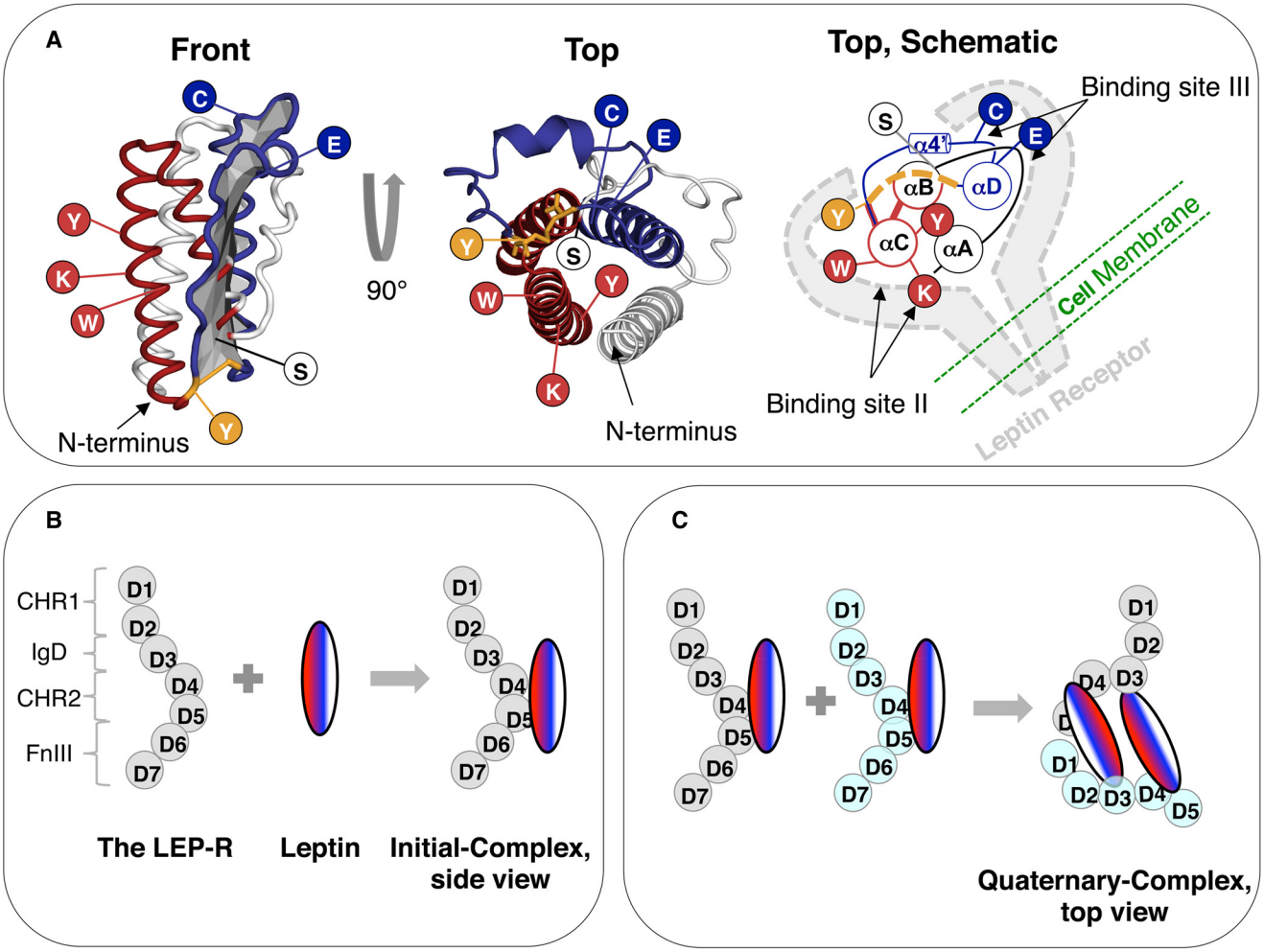

Figure 1. The PLT of leptin and the location of disease-causing mutations. Leptin has a PLT, with one disulfide bridge (*yellow*) forming a covalent loop (*deep blue*) where part of the backbone is threaded through (*deep red*). The receptor-binding sites (II in loop I and binding site III formed by the bottom part of helices A and C) are located on opposite sides of the covalent loop, necessitating the threaded topology so that both sites are "free" to interact with the receptor. The seven full-length mutations in the current work are represented as *spheres*. The *front view* of the protein is shown in the *left panel*, the *top view* in the *middle panel*, and a *cartoon representation* of the *top view* in the *right panel*. PyLasso was used to create the surface shown in the *front view representation* (70). *B*, *schematic representation* of the initial 1:1 LEP-R complex. The seven domains of LEP-R are represented as *spheres*, and leptin is *colored* as in *A*. Domains D4 and D5 can interact with receptor-binding site II in leptin (composed of Asp^{9}, Thr^{12}, Lys^{15}, Arg^{20}, Gln^{75}, Asp^{85}, and Leu^{86} (28). *C*, *schematic representation* of the 2:2 LEP-R complex. The seven domains of LEP-R are represented as *spheres* (*gray* and *cyan* to represent the two receptors interacting with leptin), and leptin is *colored* as in *A*. Once bound, two LEP-R complexes can form the active quaternary complex through interaction of D3 and binding site III on leptin. At present, it is still unclear whether receptor-binding site III is composed of residues Ser^{120}–Thr^{121}, Leu^{39}–Ile^{42}, or both (28). For clarity, domains D6 and D7 are excluded from the *top view*.

each amino acid. The size of the letter indicates the level of conservation; the larger the letter the more conserved is the residue. Five of the 13 disease-associated mutations identified in leptin are deletion or nonsense mutations expressing truncated versions of the protein: $I14_{del}$ (48, 49), $Q28T_{fs*23}$ (52), $Q34_{stop}$ (50), $G112V_{fs*15}$ (42), and $L140G_{fs*10}$ (48) (Fig. S1, *top*). Two mutations ($R84_{stop}$ (41, 43, 47) and V124E (47)) have been found in murine leptin, where $R84_{stop}$ is another truncated version, whereas V124E is a single point mutation that results in a full-length protein. All seven of the above-mentioned leptin variants except V124E are probably degraded quickly and cleared from serum due to the decrease in protein stability or inability to fold into native-like structures. Therefore, none of these mutations are considered explicitly in this study (see supporting information, "Disease-associated mutations identified in leptin").

In the current work, we study all known full-length variants of leptin that lead to leptin deficiency, namely L51S (55), D79Y (17, 19), N82K (17, 18, 46, 51), R84W (4, 44), C96Y (45), S120C (45), and V124E (47) (Figs. 1*A* and 2). Each of these mutations is systematically analyzed below.

WT leptin

Naturally occurring leptin has two tryptophan residues (Trp^{100} in loop III and Trp^{138} at the C terminus of αD). The W100E variant of leptin (pseudo-WT) was made for crystallization purposes and has since been shown to have WT-like biological activity (57). The W100E mutation is present in all of our disease-causing variants in the current work. The pseudo-WT protein forms soluble dimers/oligomers as well as visible aggregates in solution. Some of these native-like aggregates can revert to the monomeric form upon dilution or decreased tem-

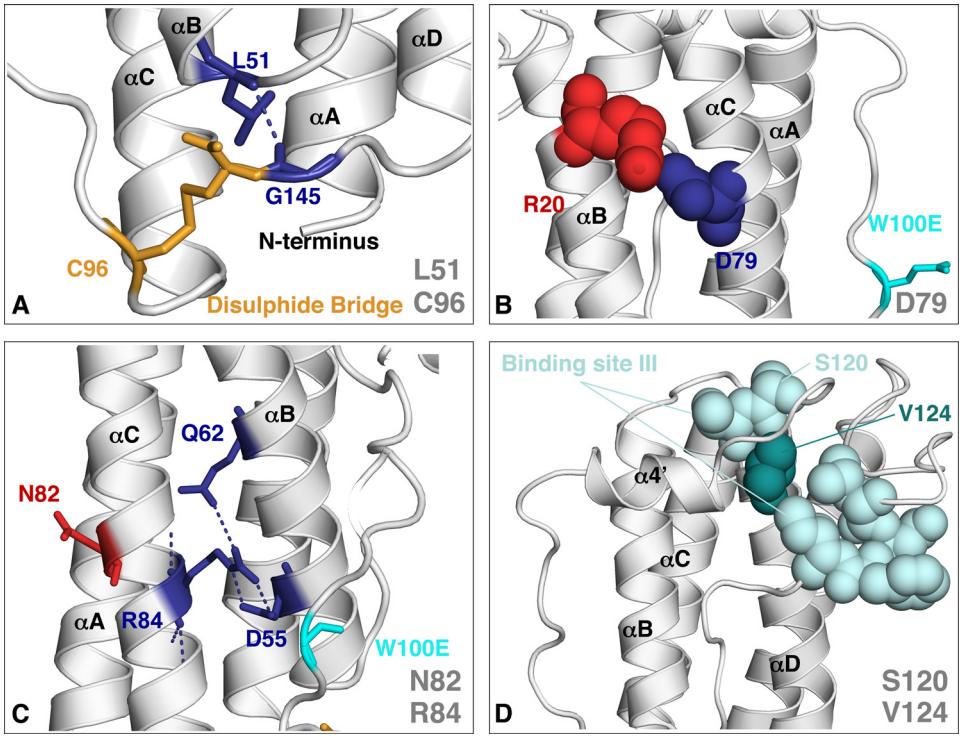

Figure 2. The structure of leptin and the location of disease-associated mutations. *A*, Leu[51] is the only buried residue that is mutated in a disease-associated full-length variant of leptin. It makes an important polar contact with Gly[145] (*deep blue dashed line*), in close proximity to the covalent loop (*yellow sticks*). This contact tethers the threaded element to the covalent loop, keeping it in place. The disulfide bridge is represented in *yellow*, indicating residue Cys[96]. *B*, Asp[79] makes a salt bridge with Arg[20] in αA. More importantly, both D79Y and N82K interact with D4 and D5 on the LEP-R. *C*, the side chain of Arg[84] is oriented toward the core of the protein interacting with Gln[62] and Asp[55], whereas Asn[82], represented as a *red stick*, is oriented toward the solvent and interacts with LEP-R. *D*, Val[124] is sandwiched in between the two proposed sites of binding site III in leptin (28). Introducing a negatively charged glutamic acid at this position could have a dramatic effect on receptor interaction. Residue Ser[120] is part of the receptor-binding site III and is located above Val[124].

perature, whereas some are irreversible aggregates that fall out of solution (58). Both types of aggregation occur during the refolding process during purification *in vitro* and are removed from the samples through size-exclusion chromatography and ultracentrifugation (see "Materials and methods").

Characterization methods for disease-associated leptin variants

To evaluate the effects of disease-associated mutations on the native structure of leptin, we conducted heteronuclear single quantum coherence (HSQC) NMR experiments (Fig. 3) together with native state dynamics (NSD) simulations for the fully folded variants using all-atom structure-based models (see "Materials and methods" for more details of the model (59, 60)). Additionally, kinetic unfolding simulations of the pseudo-WT protein were performed to characterize the threading mechanism and to study potential kinetic traps on the folding pathway using all-atom structure-based models. DSC thermoanalytical experiments were used to quantify the mutational effects on protein stability (Table 1) combined with the predictions of stability changes upon mutations using mCSM (61) to compute the predicted stability changes for mutations where a DSC signal was not observed. The prediction tool mCSM provides quantitative estimates of the distribution of stability changes induced by mutations through a novel approach to predicting the effects of mutation in proteins using graph-based signatures (61). In addition to domain stability, we calculated predicted solubility changes between WT leptin and all variants of leptin obtainable via a nonsynonymous SNP (nsSNP) using CamSol (Fig. 4) (62).

Location and biophysical character of disease-associated mutations in the structure of leptin

Disease-associated mutations in leptin are scattered throughout the structure and are not confined to one specific region (Fig. 1*A* and Fig. S1, *top*). Residue Leu[51] is located in a conserved region at the N terminus of αB (Fig. S1), in close proximity to the disulfide bridge. The aliphatic side chain is buried and orients the backbone amide for hydrogen bonding with the carbonyl of Gly[145], thereby tethering the threaded helical hairpin (αC and half of αB) to the covalent loop (Fig. 2*A*).

Residue Asp[79] is located in the middle of αC and is part of the conserved region interacting with the LEP-R through binding site II in leptin (19) (Figs. S1 and S2). The negatively charged side chain of Asp[79] forms a salt bridge with Arg[20] in αB that will break upon mutation to a noncharged tyrosine (Fig. 2*B*). Additionally, a tyrosine at this position could block the receptor from binding, as the side chain is larger and bulkier than the side chain of aspartic acid (Fig. S2). Additionally, substituting a charged side chain with a hydrophobic side chain could have dramatic effects on receptor binding.

Residue Asn[82] is located in the middle of αC and is part of a moderately conserved region interacting with the LEP-R through binding site II in leptin (Fig. S2 and Fig. 2*C*) (51). Sequence alignments to other leptin homologs show that vari-

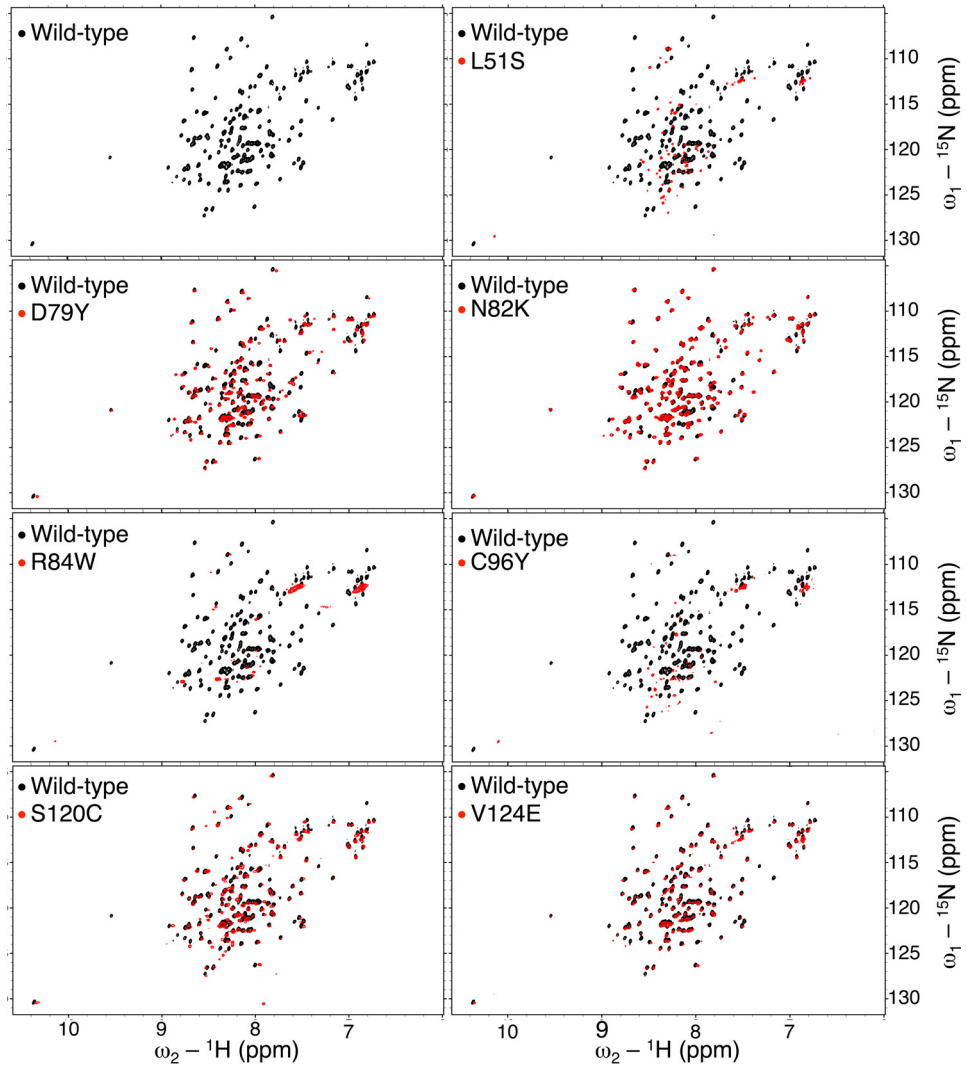

Figure 3. Structural analyses of the disease-associated mutations identified in leptin though NMR HSQCS. The ^{1}H-^{15}N HSQCs are shown for the leptin variants as a representation of the structural changes in solution at 25 °C upon mutation, where the pseudo-WT protein is represented in *black* and the overlaid mutations are represented in *red*.

Table 1
DSC experiments

Name	T_m [a]
	°C
WT (W100E)	58.7
L51S	ND[b]
D79Y	59.9
N82K	72.0
R84W	ND
C96Y	ND
S120C	71.9
V124E	58.7

[a] The raw data are provided in Fig. S3.
[b] ND, not determined.

ation at this position is permitted as long as the polarity of the side chain is kept intact (Fig. S1, *bottom*).

Residue Arg^{84} is located in the middle of αC oriented toward the core of the protein, centered in the back of the receptor-binding site II. Sequence alignments show that this region is moderately conserved (Fig. S1, *bottom*), suggesting that the mutation R84W is allowable and should not interfere with the LEP-R interaction (Fig. 2*C*). Upon mutation to tryptophan, the salt bridge formed with the side chain of Asp^{55} will break. This mutation may also affect the polar contact formed with Gln^{62} due to the introduction of a bulky side chain (Fig. 2*C*). Finally, R84W substitutes a charged, partially exposed residue with a hydrophobic residue, thereby possibly affecting the solubility of the folded state and its aggregation propensity (Table 2).

Residue Cys^{96} connects loop III in leptin with the C-terminal cysteine under oxidizing conditions (Figs. 2 and 3*A*), forming a PLT (see supporting information) (33, 35). The sequence alignments (Fig. S1, *bottom*) show that both cysteine residues are absolutely conserved (33, 39). Hence, mutating Cys^{96} to any other residue will break the disulfide bond and, consequently, disrupt the threaded topology, shifting the folding equilibrium toward a denatured state in leptin (33, 35). In this way, C96Y is analogous to the *reduced* version of the WT protein, where the disulfide bridge is chemically modified, as neither protein can form the native disulfide bridge. It has already been shown that this will lead to a detrimental loss of global stability of the protein and decreased biological activity (33, 35).

Residue Ser^{120} is located at the end of the loop between α4′ and αD, in a moderately conserved region of the protein (Fig.

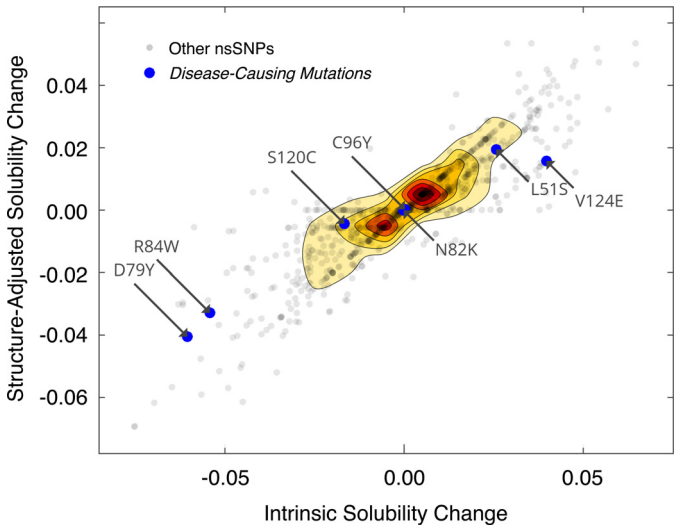

Figure 4. Predicted changes in solubility of the folded and denatured states of leptin upon mutation using CamSol (62). The predicted solubility changes for mutations in leptin are plotted against the background of all point mutations that can arise from nsSNPs. Negative values correspond to a predicted decrease in solubility. Positive values correspond to a predicted increase in the solubility compared with the WT protein. Solubility changes for all unique amino acid mutations that can arise from a nsSNP are shown as *gray dots*. Disease-causing mutations are *highlighted* with *larger blue dots*. The *density contour* indicates the density of predicted solubility changes in a particular area of the plot and is provided for clarity. A high density of mutations (*red*) is found in the regions of the plot corresponding to mutations that are predicted to slightly increase (positive values) or decrease (negative values) the solubility. Lower-density regions are indicated in *orange* and *yellow*. Both the R84W and D79Y mutations are found in the very-low-density (*white*) region of the plot, indicating that they are outliers with respect to their predicted decrease in the solubility of both the folded and unfolded states of leptin.

S1, *bottom*). It has two polar contacts: one with the backbone of Val^{123} and one with Val^{124} at the N terminus of αD, rigidifying and stabilizing loop IV (Fig. 2*D*). It has been proposed that Ser^{120} together with Thr^{121} is part of the receptor interphase III, binding to D3 of LEP-R, where binding site III is involved in the formation of the active quaternary complex (53).

Residue Val^{124} is located at the N terminus of αD in a moderately conserved region of the protein. Sequence alignments show that leucine can also be observed at this position, indicating that there is space to fit a larger side chain at this position (Fig. S1, *bottom*). However, introducing a negatively charged side chain could affect receptor-binding site III in leptin, as residue 124 is sandwiched in between the two suggested regions that interact with D3 in LEP-R (Fig. 2*D*) (28).

Structural evaluation

To evaluate the structural and dynamic perturbations of the mutations, we performed NMR HSQC experiments together with NSD *in silico* simulations for the proteins that showed fully folded spectra by NMR. The HSQCs revealed that four of the seven mutations tested fold into native-like structures exhibiting only minor shifts compared with the pseudo-WT protein, which are expected for a single point mutation (Fig. 3). The HSQC spectra of D79Y, N82K, S120C, and V124E coincide well with the pseudo-WT spectrum (Fig. 3). The spectrum of S120C shows a small population of aggregated protein, presumably due to dimerization via the introduction of a free cysteine on the surface of leptin. Native state dynamics simulations were conducted for all leptin variants that appeared to be folded according to NMR (Fig. 5). All of the folded protein variants except V124E exhibit WT-like native state dynamics in their free forms (Fig. 5*A*), indicating that the dynamic motions are not affected by these mutations. In the case of V124E, however, the mutation introduces larger fluctuations (compared with the WT protein) around residues 33–43 and 103–127 both in the free form and when bound to D4D5, thus affecting regions that are important in receptor interactions and biological activity (Fig. 5*B*).

In contrast to the well-folded variants, the HSQC backbone spectral dispersions are markedly abolished for the three remaining proteins, with only between ~1.3 and 0.98 ppm over ^{1}H and between ~18 and 20 ppm over ^{15}N dimensions. R84W display the most dramatic shifts, indicating a nearly complete denaturation or an aggregated protein (Fig. 3). The HSQCs of L51S and C96Y show a spectrum typical of a protein that exhibits partially denatured/partially folded states (Fig. 3), where the main population seems to be denatured. There are, however, some peaks that overlap with the pseudo-WT protein, indicating that there might be a small population of folded protein in these samples. However, higher-resolution NMR data are necessary to make further conclusions regarding the native structures of these proteins.

Biophysical characterization of stability

DSC measurements were performed to evaluate the mutational effects on protein stability. The four variants that appeared to have native-like structures by NMR (D79Y, N82K, S120S, and V124E) have apparent melting temperatures the same as or higher than that of the pseudo-WT protein (Table 1). The other leptin variants (L51S, C96Y, and R84W), which showed poorly dispersed NMR spectra, do not exhibit an unfolding transition under any conditions that we tested. This suggests that L51S, C96Y, and R84W are denatured and/or aggregated at room temperature. Because the C96Y variant is analogous to the reduced form of WT leptin, where the disulfide bridge is chemically removed, we expect that the loss of the disulfide bridge through mutation has a similar effect on protein stability as reducing the pseudo-WT protein, which results in a destabilization of $\Delta\Delta G = -1.6$ kcal/mol (33). As there are no clear unfolding signals from DSC experiments for L51S and R84W either, we used the web tool mCSM to estimate the stability changes upon mutation (61). L51S is predicted to be significantly destabilized, with a predicted $\Delta\Delta G$ of −2.882 kcal/mol, whereas R84W is predicted to be only marginally destabilized ($\Delta\Delta G$ of −0.649 kcal/mol).

Aggregation during purification

Leptin is purified from inclusion bodies *in vitro*, which involves initially denaturing the protein in its reduced state. After the first purification step, both *reducing* and *oxidizing* GSH are used to allow the disulfide bridge to be made and broken upon refolding, increasing the probability of successful threading of the covalent loop or refolding without the disulfide bridge. Ultracentrifugation and an extra size-exclusion step are used to eliminate aggregated and/or oligomeric species. In the

Table 2
Biophysical properties of the disease-associated mutations identified in leptin

Name	pI[a]	Net charge at pH 7.0	Original amino acid	Mutated amino acid	Sequence conservation[b]	Position
WT	6.16	−2.2				
pswt (W100E)	5.76	−3.2	Hydrophobic	Negative	Low	Loop III
L51S	5.76	−3.2	Hydrophobic	Polar	High	αB-loop II
D79Y	6.16	−2.2	Negative	Hydrophobic	High	αC
N82K[c]	6.16	−2.2	Polar	Positive	Moderate	αC
R84W[d]	5.41	−4.2	Positive	Hydrophobic	Moderate	αC
C96Y	5.78	−3.1	Special	Hydrophobic	High	Loop III
S120C[e]	5.76	−3.2	Polar	Special	Moderate	Loop IV
V124E[f]	5.76	−3.2	Hydrophobic	Negative	Moderate	αD

[a] A theoretical pI calculated for each mutation using ExPASy Bioinformatics Resources Portal (https://web.expasy.org/compute_pi/).
[b] Supported by the sequence alignment in Fig. S1 (*bottom*).
[c] Residue 82 is usually an asparagine but can also be a threonine (polar), as is found in some leptin homologs.
[d] Residue 84 is usually an arginine but can also be a leucine (hydrophobic), as is found in some leptin homologs, indicating that a hydrophobic tryptophan residue should not cause disease, as the size agrees with an arginine and the sequence allows for a hydrophobic substitution. Nonetheless, a tryptophan residue at this position may block threading through the covalent loop due to its large "sticky" side chain that can interact with the other tryptophans in the covalent loop (*i.e.* Trp^{100} and Trp^{138} in the WT protein.
[e] Residue 120 is usually a serine but can also be a threonine (polar), as is found in some leptin homologs. Adding a reactive cysteine to the protein might lead to dimerization, as residue 120 is at the surface and thus free to interact with other leptin molecules to form an intermolecular disulfide bond under *oxidizing* conditions in the cell.
[f] Residue 124 is usually a valine but can also be a leucine (hydrophobic), as is found in some leptin homologs.

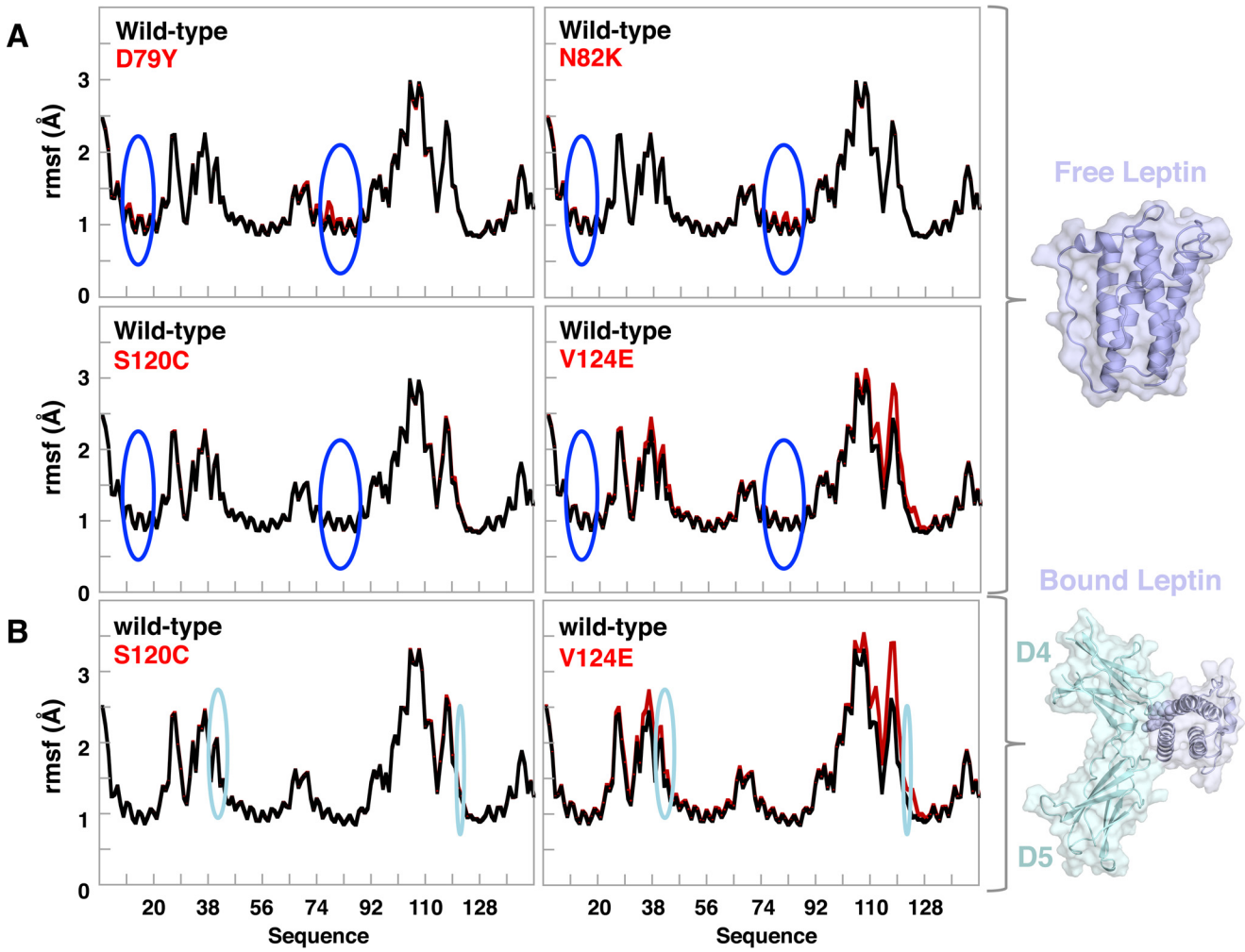

Figure 5. The backbone RMSF of residues in leptin is plotted against the amino acid sequence index for several leptin variants. The WT protein is represented in *black*, and the disease-associated variants are shown in *red*. *A* and *B*, plots of RMSF for the free protein (*A*) and protein bound to D4 and D5 (*B*) (32). The receptor interaction sites are *highlighted* with *ellipses*, where receptor-binding site II is represented in *blue* and binding site III in *cyan*.

case of R84W, no protein was obtained using this standard method. Therefore, *reducing* agent alone was used to purify this protein. The failure to obtain R84W via the standard purification indicates that this particular variant cannot refold at all in its *oxidized* state, whereas in the *reduced* state, purification of some protein is possible. The different behavior of the reduced and oxidized forms of R84W points to an involvement of the PLT in the misfolding of the oxidized state, where threading of the covalent loop must occur in order to reach the native topology. Arg^{84} is part of the threaded element, and substitution by

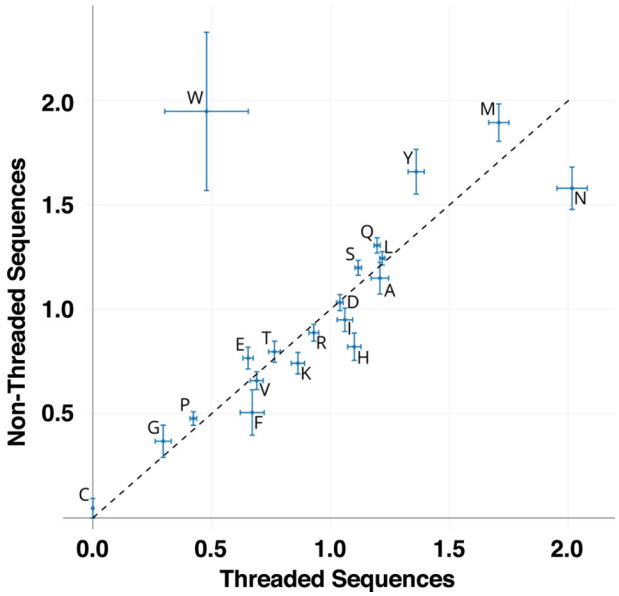

Figure 6. Relative prevalence of amino acid types in the threaded element for threaded and nonthreaded leptin homologs. In threaded sequences, tryptophan shows up about half as often in the threaded element as it does in the sequence as a whole. In nonthreaded sequences (defined as sequences that lack one or both cysteines at positions 96 and 146), tryptophan appears about twice as much in the part of the sequence that corresponds to the threaded element as it does in the sequence as a whole. Across all amino acid types, this difference in relative prevalence between threaded and nonthreaded sequences is largest for tryptophan. The mean values and S.D. (shown as *error bars*) were obtained by bootstrapping (see "Materials and methods" for details).

tryptophan may result in steric clashes during loop threading. MD simulations with an all-atom structure-based model indicate that Arg^{84} forms transient nonnative interactions during threading with Trp^{138} as well as W100E (Fig. 6). When position 84 is mutated to tryptophan, these nonnative interactions would be stabilized and might trap the protein in a partially threaded configuration. Taken together, it seems plausible that the extreme aggregation propensity of R84W under *oxidizing* conditions is due to unsuccessful threading through the closed covalent loop. Additionally, having tryptophan in the threaded element compared with the sequence as a whole is disfavored in leptin homologs only when the PLT is formed (see Fig. 6 and "Materials and methods"). This bioinformatic analysis serves as further evidence that having a tryptophan in the threaded element may prevent or slow down threading, which is consistent with our experience when trying to purify the R84W variant in reducing and oxidizing conditions.

For R84W refolded under *reducing* conditions, our HSQC spectrum also shows little dispersion, and most of the native peaks are missing. The low dispersion in NMR and the lack of an observable DSC transition, together with previously published results (4, 44), could indicate that R84W, even when prepared under reducing conditions, is not natively folded but denatured or aggregated. We used the CamSol algorithm to investigate solubility changes upon mutation in leptin (62). R84W is predicted to be much less soluble than the WT, with a higher propensity to aggregate even from its folded state (Fig. 4), consistent with the observations of aggregation from the experiments. Although the R84W mutation is predicted to be only modestly destabilizing to the native structure by mCSM, mCSM does not take into account the possibility that a tryptophan at position 84 may prevent threading of the covalent loop and thereby block access to the native PLT conformation. Whether the R84W mutation contributes to aggregation either by preventing threading, reducing the solubility of the folded state, or both, remains to be definitively determined. The currently available data suggest that both mechanisms are possible. Interestingly, CamSol also predicts that the D79Y mutation significantly reduces the solubility of leptin (Fig. 4). We have not observed any indication of aggregation for D79Y over the time scale of our experiments. However, this does not preclude the possibility of aggregation on biologically relevant time scales.

Discussion

The World Health Organization has designated obesity a global epidemic, and the underlying mechanisms are not yet completely understood. Obesity is the result of a combination of social, environmental, and genetic factors. It leads to metabolic syndrome, depression, and type-2 diabetes (63). Research into the genetics of human obesity has become increasingly sophisticated with respect to molecular technologies, biostatistics, and efficient design strategies. However, the molecular mechanisms behind genetic obesity and leptin deficiency remain elusive.

To date, 13 disease-associated mutations in leptin have been identified in mice and humans (Fig. S1) (18, 19, 41–52). In the current work, we investigate structural perturbations and the importance of the PLT in mutations leading to congenital leptin deficiency. Our results, considered in the light of the currently available literature, suggest that the disease-associated mutations in leptin may lead to leptin deficiency by at least four different mechanisms: (i) blockage of receptor interface through binding site II in leptin, interacting with D4 and D5 of the LEP-R initiating a nonsignaling complex; (ii) modifications of the affinity for D3 on the LEP-R through binding site III in leptin, perturbing the formation of the quaternary complex and leading to impaired signaling; (iii) destabilization of the native state through the loss of salt bridges and/or the disulfide bridge leading to an unthreaded topology, probably leading to degradation/clearance from serum; and (iv) unsuccessful threading through the covalent loop that might lead to protein misfolding and/or a loss of solubility of the folded leptin structure.

We hope that these classifications will serve as a framework for understanding the possible ways in which mutations in leptin can give rise to leptin deficiency. To our knowledge, no such framework has yet been proposed. This framework should be revised when more structural, biological, and biophysical data of leptin variants become available.

Mutations affecting receptor-binding sites II and III in leptin, through mechanism (i) or (ii)

Our NMR experiments show that four of the seven mutations in the current work fold into WT-like structures with maintained or higher thermal stability, namely D79Y, N82K, S120C, and V124E (Fig. 3 and Table 1). Interestingly, inspection of the LEP-R complex together with previously published results indicate that all of these leptin variants are involved in receptor binding (28). Niv-Spector *et al.* (18) showed that Asp^{79}

and Asn^{82} are part of receptor-binding site II, forming the interface with D4D5 (18), whereas Ser^{120} together with Thr^{121} has been proposed to form part the binding interface of D3 (18, 53, 64). When mutated, D79Y and N82K leptin variants fail to dock with the LEP-R, acting as a leptin antagonist blocking leptin signaling (Fig. S2) (16). Both D79Y and N82K have been studied in human cells and were shown to have a decreased affinity for the LEP-R and cannot initiate signaling for this reason (18, 19).

In contrast, for V124E, Hong *et al.* (47) performed co-immunoprecipitation analysis showing that the variant was indeed interacting with the LEP-R in hypothalamus through receptor-binding site II. However, to our knowledge, no further studies investigating the second step of receptor binding and signaling for V124E have been conducted. Indeed, both S120C and V124E have the same affinity for the LEP-R as the WT protein without being able to initiate leptin signaling in the cells (45, 47). This suggests that neither of these two mutations affects receptor-binding site II in leptin and the formation of the initial nonsignaling complex.

The locations of Ser^{120} and Val^{124} further suggest that both S120C and V124E interfere with the subsequent formation of the active quaternary complex (Fig. 2*D*). In the case of V124E, our MD simulations support this conclusion further by showing increased fluctuations in binding site III, both for the free V124E variant and bound to D4D5 of LEP-R in a model of the initial complex (Fig. 5). As residue Val^{124} is sandwiched between the two proposed regions forming binding site III in leptin, it is plausible that substitution of a larger glutamic acid residue will perturb the local dynamics. The additional effect of introducing a negative charge is not considered in our all-atom structure-based model but can also be expected to be disruptive. Additionally, Mancour *et al.* (26) showed that mutated LEP-R (*i.e.* where they introduced a mutation in residue 370 (L370A) in D3) abolishes leptin binding via leptin receptor-binding site III. Thus, we infer that perturbation of receptor-binding site III on leptin and/or perturbation of D3 on the LEP-R leads to an incomplete formation of the 2:2 quaternary complex.

According to our NMR and DSC experiments, D79Y, N82K, S120C, and V124E fold into native-like structures with comparable or increased thermal stability compared with the pseudo-WT protein (Fig. 3 and Table 1). These results agree with the previously published observations that D79Y, N82K, and V124E are biologically inactive proteins detected in serum of patients who nonetheless suffer from leptin deficiency (17, 18, 47, 51). Interestingly, no protein has been detected in serum of patients with the S120C mutation (45, 47). The impaired secretion in the case of S120C could be caused by the added reactive cysteine forming intermolecular disulfide bridges in the cell under *oxidizing* conditions. It is also possible that S120C is secreted, but no protein is found in serum of patients depending on the immunoassay used for leptin detection, as was shown in the case of N82K (17–19, 46, 51, 65). In the case of N82K, the conventional assay for immunoreactive leptin (a radioactive immunoassay and ELISA and/or Western blotting) failed to detect leptin in serum of patients (46, 65), whereas the new immunoassay developed by Kratzsch *et al.* (66) is able to recognize the functionally relevant receptor-binding complex with leptin.

Taken together, we suggest that proteopathies induced by these four leptin variants (D79Y, N82K, S120C, and V124E) arise due to alterations in the binding pocket of binding sites II and II in leptin according to mechanism (i) disrupting the formation of the initial 1:1 nonsignaling complex and/or mechanism (ii) disrupting the formation of the 2:2 active quaternary complex (Fig. 8). In the case of S120C, it seems possible that this functional disruption is combined with a separate depletion of serum levels of the inactive protein.

Mutations destabilizing or misfolding the protein through mechanism (iii) and/or (iv)

None of the leptin variants containing the mutation L51S, C96Y, or R84W exhibit a thermal unfolding transition from a native to a denatured protein in our DSC measurements or show a native-like HSQC spectrum. This indicates that these proteins are denatured, aggregated, or both. Our bioinformatic studies show that both the L51S and C96Y mutations decrease the global stability of leptin compared with the pseudo-WT protein, thus shifting the folding equilibrium toward the denatured state (Figs. 4 and 5). According to the CamSol algorithm (Fig. 4), L51S and C96Y are soluble *if folded*, but the lowered stability of the native state is sufficient to explain their absence from serum. R84W, on the other hand, is predicted to maintain WT-like stability ($\Delta\Delta G$ of -0.649 kcal/mol (61)) but is also predicted by CamSol to be less soluble due to interactions introduced by the partially surface-exposed tryptophan. Interestingly, an arginine-to-tryptophan substitution is one of the most common disease-causing mutations (67).

In leptin, the effect of the mutation R84W could depend on the added complexity of threading a tryptophan though the covalent loop under oxidizing conditions, as no protein is obtained from the standard method of purification. Only when refolded under *reducing* conditions is a small concentration of protein produced during purification, suggesting that the linear (unthreaded) form of R84W can fold. Yet the protein obtained via purification in the reduced state also aggregates once the disulfide bridge is closed under the oxidizing conditions used in the NMR experiments (Fig. 3), consistent with CamSol's prediction of reduced solubility in the folded state (Fig. 4). Upon purification with both *reducing* and *oxidizing* GSH, severe aggregation occurs, indicating further impaired refolding of *oxidized* R84W when the covalent loop is closed by the disulfide bridge. Steric hindrance during threading by the bulky tryptophan residue added to the threaded element may prevent successful folding. According to our MD simulations of pseudo-WT leptin (Fig. 7), residue 84 forms nonnative interactions with the other tryptophan, Trp^{138}, located on the covalent loop. With a tryptophan introduced in position 84, these interactions threaten to trap the protein in a partially threaded misfolded configuration. This effect could be even more pronounced *in vivo*, as leptin naturally has another tryptophan residue in the covalent loop, Trp^{100}, which had been substituted in the pseudo-WT to help crystallization. This observed misfolding, in conjunction with the formation of aggregates as the protein is refolded in the *reduced* state and the disulfide bridge

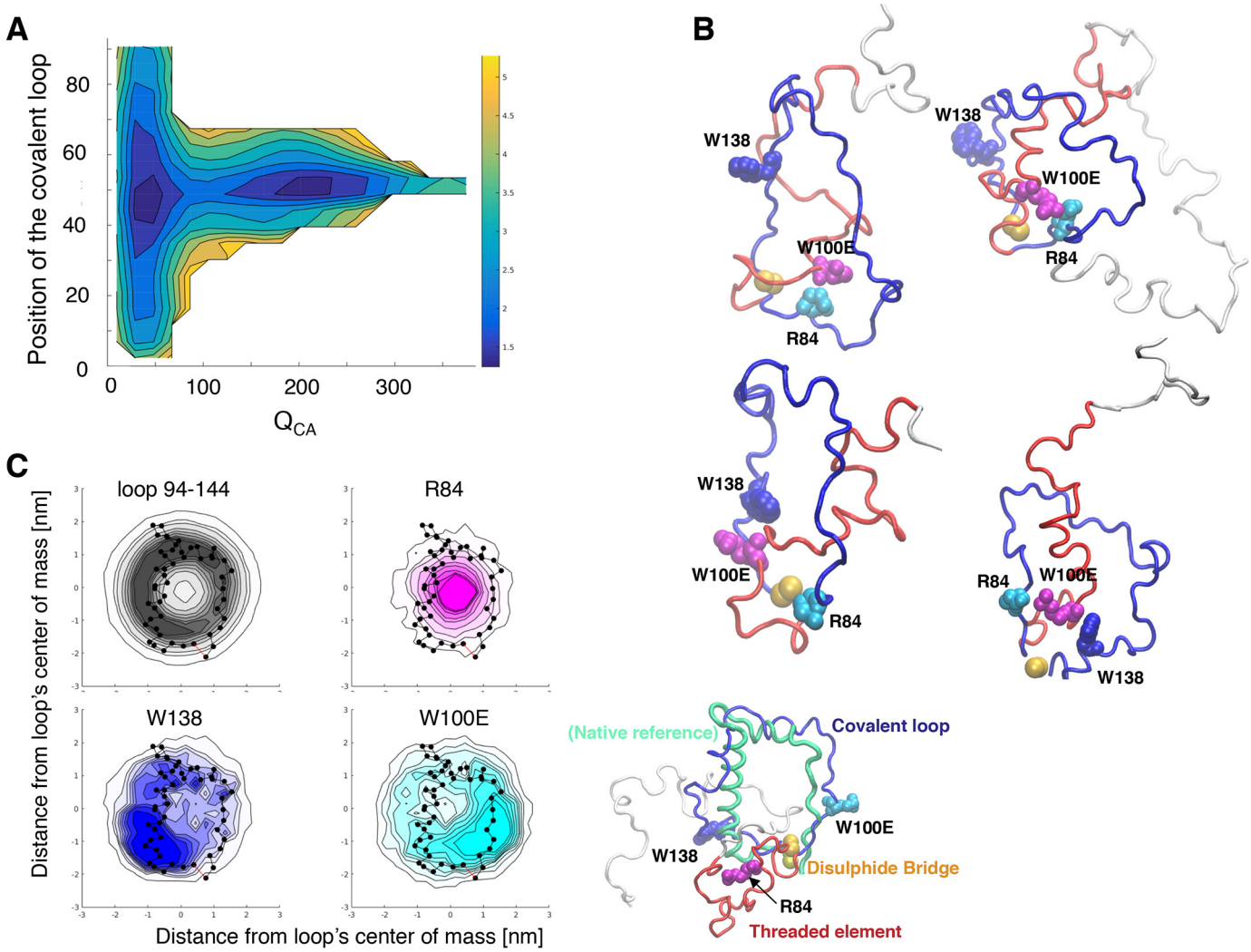

Figure 7. MD simulation of an all-atom structure-based model for the pseudo-WT of leptin (W100E). The covalent loop is represented in *deep blue*, the disulfide bridge in *yellow*, the threaded element (*i.e.* the part of the backbone that has to thread through the covalent loop to reach the native state) in *red*, and the N-terminal portion outside of the covalent loop (*i.e.* residues 1–51, which does not thread through the covalent loop) in *white*. *A*, a pseudo-free energy landscape for threading and folding, plotted as a function of the number of formed native contacts and of the index of the threaded residue, placed in the plane of the covalent loop. *B*, example structures during the unthreading event, with residue Arg^{84} (*magenta*) near the plane of the covalent loop and coming into close proximity of loop residues W100E (*blue*) or Trp^{138} (*cyan*). *C*, histograms for the positions of selected parts of the structure in the plane of the covalent loop, during unthreading via a slipknotting mechanism, with a residue from 80 to 89 in the plane of the covalent loop. Frames were aligned by fitting the covalent loop to the native configuration. The principal axes of the native covalent loop are used as coordinates. Histograms are shown for loop residues 96–146, for residue Arg^{84} in the threaded element, and for loop residues W100E and Trp^{138}. These histograms show that Arg^{84} passes in principle through the opening at the center of the loop, which appears circular on average because it is unstructured during threading. But the distributions for Trp^{138} and W100E also indicate that each may approach Arg^{84} closely in the process. The selected structures with short distances between the side chains of these pairs illustrate the possibility of such interactions.

is *reoxidized* (Fig. 8), could contribute to leptin deficiency in patients carrying the R84W mutation.

Taken together, the inability to fold, due to a loss in protein stability, in the case of L51S and C96Y explains why there is no leptin observed in serum of patients, as denatured proteins are degraded and not secreted from the cell (4, 55). R84W finally shows so many defects that it is hard to attribute its pathological effects to a single cause. But there is evidence that failure to thread into the PLT topology under oxidizing conditions is contributing to the severe aggregation observed in R84W.

Conclusions

We propose that mutations identified in patients with congenital leptin deficiency can cause proteinopathies through at least four different mechanisms. Some of the full-length disease-associated leptin variants studied here can fold into stable, native-like structures and still lead to disease through impaired function, as seen in Fig. 8. Existing data about receptor binding and our own simulation results indicate that signaling may fail either at the stage of the initial binary complex between leptin and LEP-R or at the stage of the active quaternary complex formed from two copies of the binary complex. As the currently available published data on LEP-R binding have focused on formation of the initial complex, these studies may not have picked up on deficiencies that are only apparent at the latter stage of receptor interaction and the formation of the quaternary complex and therefore report WT-like affinities for some variants. Additional

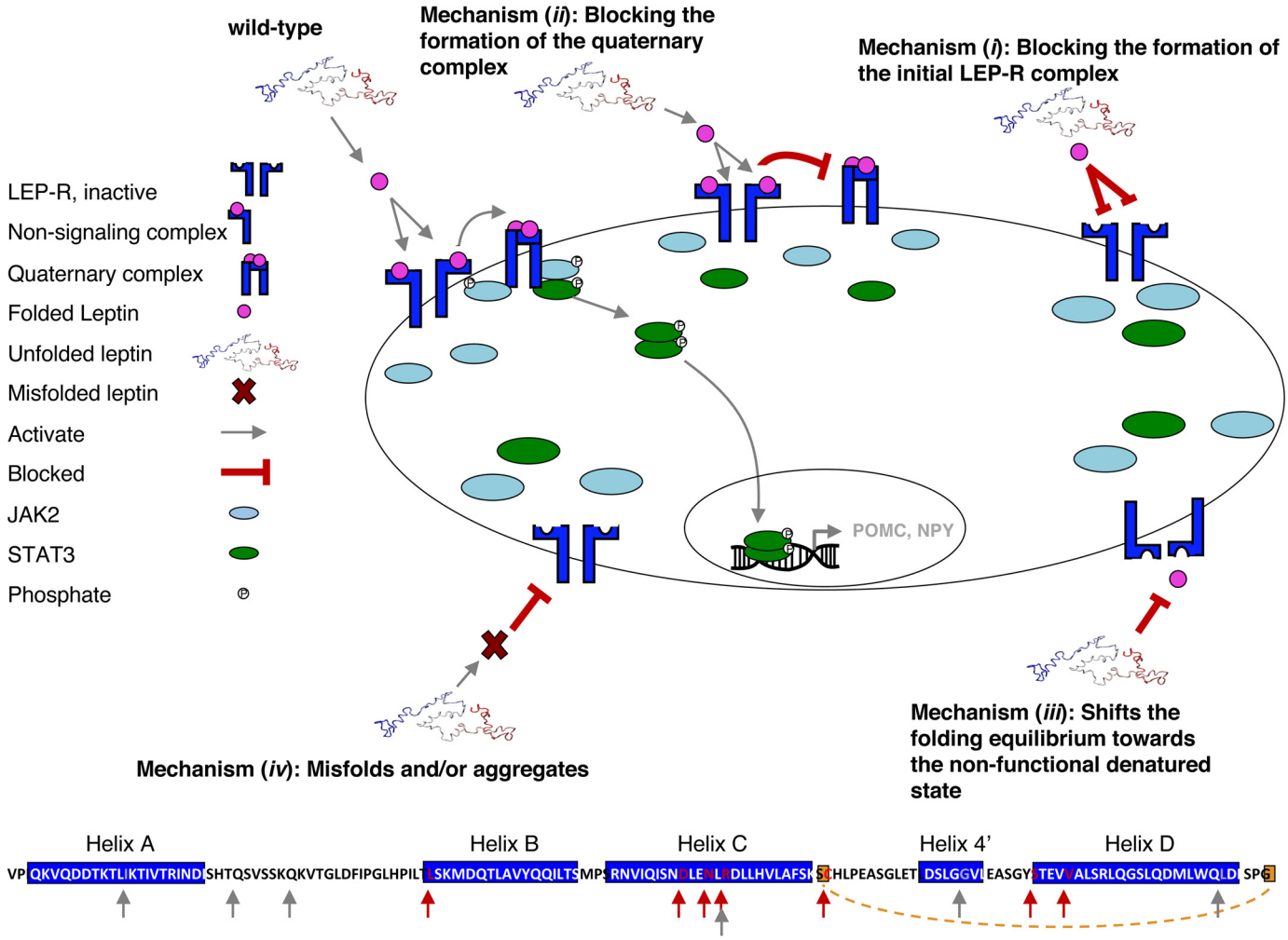

Figure 8. Schematic representation of the proposed underlying mechanisms causing the obese phenotype in patients suffering from congenital leptin deficiency. The current work suggests that the pathology from known disease mutations probably does not only occur from misfolding and/or aggregation of the protein, as was proposed earlier. Instead, we propose that leptin deficiency can arise from at least four different mechanisms, including (i) blockage of receptor interface through binding site II in leptin, interacting with D4 and D5 of the LEP-R initiating a nonsignaling complex; (ii) modifications of the affinity for D3 on the LEP-R through binding site III in leptin, perturbing the formation of the active quaternary complex and leading to impaired signaling; (iii) destabilization of the native state through the loss of salt bridges and/or the disulfide bridge, leading to an unthreaded topology, probably leading to degradation and clearance from serum; and (iv) unsuccessful threading through the covalent loop that might lead to protein misfolding and/or an aggregation, making the protein insoluble. The *bottom* of the figure shows the sequence of leptin, indicating the positions of all known disease mutations found in leptin. The full-length mutations used in the current work are represented with *red arrows*, whereas the other mutations, producing truncated versions of the protein, are represented with *gray arrows*.

studies of the later stages of signaling would be valuable to test our interpretation.

Some leptin variants lose function due to a loss of stability of the native state, misfolding, and aggregation. The resulting phenotypes correspond to classical leptin deficiency, wherein patients lack detectable levels of leptin in serum. There is clear support for loss of stability in two cases (L51S and C96Y). Because one of them (C96Y) involves loss of the covalent disulfide bridge, functional impairment analogous to *reduced* WT protein is expected as an additional consequence, but the loss of stability presumably takes effect before the defective binding can be realized. The last mutation (R84W) produces such severe aggregation under different conditions that a multiplicity of causes may contribute to loss of function. Our unfolding simulations indicate that misfolding due to disrupted threading of the covalent loop of the PLT is a possible factor.

In summary, our data, considered in the light of available literature on leptin variants, support at least four different mechanisms of disease causation by mutations in leptin (Fig. 8). The disulfide bridge is relevant because it is integral to leptin's complex pierced lasso topology, but its disruption is certainly not the only possible cause of leptin deficiency.

Materials and methods

Molecular biology and protein purification

Pseudo-WT leptin was cloned into a pET3a vector and transformed into competent BL21 *Escherichia coli* cells. Protein expression was induced by the addition of 1 mM isopropyl 1-thio-β-D-galactopyranoside, and cells were harvested and purified from inclusion bodies as described (33). When purifying R84W, one can only use *reducing* agent to produce a very small amount of protein. Normally, we use both *reduced* and

oxidized GSH during the refolding process, allowing the disulfide bridge to be made and broken. This increases the probability of successful threading of the covalent loop or folding without the disulfide bridge. If both *reducing* and *oxidizing* agents are used, no soluble protein will be observed as R84W visibly aggregates upon refolding.

All proteins were run through a size-exclusion column after refolding to eliminate oligomeric/misfolded species. Protein purity was determined by SDS-PAGE, and fractions containing the monomeric protein were collected. Disease mutations were constructed by QuikChange mutagenesis, and primers were ordered from Stratagene.

NMR HSQC experiments

^{1}H-^{15}N HSQC and HSQC TROSY experiments were collected on a 600-MHz Bruker magnet equipped with a CryoProbe at 25 °C, using a 50 mM BisTris buffer at pH 6.3 and 10% D_2O. The final protein concentration was 100 μM for all proteins except C96Y. Due to its propensity to form intermolecular disulfide bonds, 20 mM TCEP was added to the buffer used in the C96Y experiments, and the concentration was maintained at 10 μM. To avoid protein aggregation and/or degradation after purification, all NMR experiments were conducted immediately after size-exclusion chromatography and buffer exchange of the leptin variants.

Both HSQC and HSQC-TROSY experiments were collected at different temperatures and different protein concentrations for proteins that did not display native-like spectra at 25 °C. S120C was measured under oxidizing conditions. An overlay of the HSQCs collected at 25 °C is represented in Fig. 3.

DSC thermoanalytical experiments

The advantage of using a DSC assay over chemical denaturation in this study is the reduced incubation time and faster measurements. DSC also uses much smaller sample sizes at lower protein concentrations, which is important in this project, as many of the mutated proteins are hard to purify in significant amounts due to their propensity to aggregate. To avoid misinterpretations of disassociation of dimeric/oligomeric species as unfolding transitions, we plotted the T_m for the lowest protein concentration tested (0.1 mg/ml; Table 1 and Fig. S3)

All measurements were collected under *oxidizing* conditions, except in the case of C96Y, as leptin is secreted and circulates in blood under *oxidizing* conditions. A Malvern MicroCal VP-Capillary DSC system equipped with a 130-μl capillary reaction cell (Tantalum) was used to collect the thermal denaturation transition for all samples. 400 μl of sample was pipetted into a 96-well plate and tested with a scan rate of 60 °C/h with a temperature range from 20 to 90 °C using a 10-s filter. Protein concentrations tested were between 0.1 and 0.5 mg/ml in a 20 mM MES buffer at pH 6.3.

In silico conformational dynamics

All-atom structure-based simulations (59, 60) were performed to characterize the NSD of mutations demonstrating a native-like fold in the current work, both in their free and bound forms. The bound structure of the WT protein docked to D4 and D5 of the receptor was taken from Carpenter *et al.* (Protein Data Bank code 3V6O (32)), and leptin mutations were generated in PyMOL. Simulation inputs, containing all individual heavy atoms, were prepared from the structures using the SMOG version 2 (68) package, with a customized template to account for the disulfide link. A covalent bond, represented by a harmonic potential, was placed between the two cysteine sulfur atoms, whereas tertiary interactions between the two cysteine residues were dropped. To capture the assumed locally destabilizing effects of the mutations, tertiary interactions formed by the respective mutated residues were deleted. All other native tertiary interactions are represented by stabilizing contact potentials in the model, in accordance with the principle of minimal frustration. All systems were simulated in the native state at a reduced temperature of $T/T^* = 0.83$, using an adapted version of the GROMACS software package. (The reference temperature T^* is close to the folding temperature due to the construction of the model.) Four independent runs were performed for each case, each for 120 ns of simulation time for the unbound proteins and for 280 ns for the receptor-bound ones. Dynamics in the native state were characterized as root mean square fluctuation (RMSF) around the mean configuration. All discussed mutation effects are much larger than the *error bars*. For the pseudo-WT protein, 180 additional unfolding runs were performed, each over 10 ns of simulation time at a temperature of $T/T^* = 1.17$, which is above the folding temperature of the system. All 180 runs unfolded and unthreaded during the simulations, and no significant (re)threading occurred. The mechanism was characterized by computing the number of native contacts at every step and by using the PyLasso software to determine the degree of unthreading in each configuration. A pseudo-free energy landscape plotted for these two coordinates indicates that unfolding precedes unthreading, which finally occurs in the context of an unstructured chain with a low number of native contacts. 147 runs (82%) unthreaded via a slipknotting mechanism, which is the dominating pathway for leptin (33, 35). Plugging of the covalent loop was found in the remaining cases. To study nonnative interactions formed by residue Arg^{84} during unthreading, first those frames were selected from the slipknotting runs where a residue near Arg^{84} was placed in the plane of the covalent loop according to PyLasso. All of these structures were then aligned by fitting their loop configuration to the native structure of the covalent loop. Histograms for the placement of residues of interest in the plane of the loop were collected, using the principal axes of the native covalent loop as coordinates.

CamSol-predicted changes in solubility

The CamSol method (68) was used to predict changes in solubility compared with WT leptin. Unlike the structurally corrected scores described previously (68), the structure-adjusted scores that we are showing here have been obtained by setting all of the "smoothing weights" (68) to 1 and by using the solvent exposure of the WT residues for the "exposure weights." Therefore, these structure-adjusted scores are corrected for solvent exposure in the WT structure but not proximity in three-dimensional space. The change in intrinsic and structure-adjusted CamSol score upon mutation is plotted for

every point mutation that can be reached from the WT leptin sequence via a nonsynonymous SNP.

Multiple-sequence alignment of leptin homologs

Analysis of the relative prevalence of amino acid types in the threaded element of leptin homologs for threaded and nonthreaded sequences was performed as follows. We obtained the multiple-sequence alignment of leptin homologs from PFAM (69) and divided the sequences into threaded and nonthreaded sequences by looking at the positions corresponding to Cys^{96} and Cys^{146} in WT leptin. Sequences with cysteines in both of these positions were considered threaded, and all other sequences were considered nonthreaded. Note that sequences in the alignment that do not have cysteines at those positions also do not have cysteines in nearby positions, so the classification of threaded and nonthreaded sequences would not depend on subtle changes to the alignment. Of a total of 279 sequences in the PFAM alignment of leptin homologs (PFAM ID PF02024, PFAM version 31.0), using the above definition, 192 are threaded and 87 are nonthreaded. Separately for the threaded and nonthreaded sets of sequences, we created 1000 new alignments of equal size (192 for threaded sequences and 87 for nonthreaded sequences) by drawing randomly from the available sequences with replacement. For each alignment, we then computed the frequency of the appearance of each amino acid type in the threaded element and in the sequence as a whole. The values reported in Fig. 6 are the ratios of these two values for each amino acid type in threaded and nonthreaded sequences. The mean (*points*) and S.D. (*error bars*) of the distributions of values obtained by computing this ratio for each amino acid type on 1000 alignments are plotted in Fig. 6. Gaps were excluded from all frequency calculations.

Author contributions—E. H. conceptualization; E. H. and L. N. formal analysis; E. H. validation; E. H. investigation; E. H. writing-original draft; E. H., L. N., N. P. S., H. L., P. A. J., and J. N. O. writing-review and editing; N. P. S. and H. L. data curation; P. A. J. and J. N. O. funding acquisition.

Acknowledgments—We thank Vivien Chau for help with cloning, Xuemei Huang for help with NMR experiments, and Verna Frasca and Muneera Beach from Malvern for all of the help with the DSC measurements and analysis. We thank Pietro Sormanni for running the CamSol calculations. Finally, we thank Jens Danielsson and Ryan R. Cheng for helpful scientific discussions.

References

1. Ramachandrappa, S., and Farooqi, I. S. (2011) Genetic approaches to understanding human obesity. *J. Clin. Invest.* **121,** 2080–2086 CrossRef Medline
2. Gibson, W. T., Farooqi, I. S., Moreau, M., DePaoli, A. M., Lawrence, E., O'Rahilly, S., and Trussell, R. A. (2004) Congenital leptin deficiency due to homozygosity for the Δ133G mutation: report of another case and evaluation of response to four years of leptin therapy. *J. Clin. Endocrinol. Metab.* **89,** 4821–4826 CrossRef Medline
3. Ozata, M., Ozdemir, I. C., and Licinio, J. (1999) Human leptin deficiency caused by a missense mutation: multiple endocrine defects, decreased sympathetic tone, and immune system dysfunction indicate new targets for leptin action, greater central than peripheral resistance to the effects of leptin, and spontaneous correction of leptin-mediated defects. *J. Clin. Endocrinol. Metab.* **84,** 3686–3695 CrossRef Medline
4. Strobel, A., Issad, T., Camoin, L., Ozata, M., and Strosberg, A. D. (1998) A leptin missense mutation associated with hypogonadism and morbid obesity. *Nat. Genet.* **18,** 213–215 CrossRef Medline
5. Bray, G. A., and York, D. A. (1979) Hypothalamic and genetic obesity in experimental animals: an autonomic and endocrine hypothesis. *Physiol. Rev.* **59,** 719–809 CrossRef Medline
6. Elias, C. F., and Purohit, D. (2013) Leptin signaling and circuits in puberty and fertility. *Cell Mol. Life Sci.* **70,** 841–862 CrossRef Medline
7. Garofalo, C., and Surmacz, E. (2006) Leptin and cancer. *J. Cell. Physiol.* **207,** 12–22 CrossRef Medline
8. Fietta, P. (2005) Focus on leptin, a pleiotropic hormone. *Minerva Med.* **96,** 65–75 Medline
9. Margetic, S., Gazzola, C., Pegg, G. G., and Hill, R. A. (2002) Leptin: a review of its peripheral actions and interactions. *Int. J. Obes. Relat. Metab. Disord.* **26,** 1407–1433 CrossRef Medline
10. Poirier, P., Giles, T. D., Bray, G. A., Hong, Y., Stern, J. S., Pi-Sunyer, F. X., and Eckel, R. H. (2006) Obesity and cardiovascular disease: pathophysiology, evaluation, and effect of weight loss. *Arterioscler. Thromb. Vasc. Biol.* **26,** 968–976 CrossRef Medline
11. Leibel, R. L., Bahary, N., and Friedman, J. M. (1990) Genetic variation and nutrition in obesity: approaches to the molecular genetics of obesity. *World Rev. Nutr. Diet.* **63,** 90–101 CrossRef Medline
12. Leibel, R. L., Bahary, N., and Friedman, J. M. (1993) Strategies for the molecular genetic analysis of obesity in humans. *Crit. Rev. Food Sci. Nutr.* **33,** 351–358 CrossRef Medline
13. Dickie, M. M., and Lane, P. W. (1957) Plus letter to Roy Robinson 7/7/70. *Mouse News Lett.* **17,** 52
14. Farooqi, I. S., Matarese, G., Lord, G. M., Keogh, J. M., Lawrence, E., Agwu, C., Sanna, V., Jebb, S. A., Perna, F., Fontana, S., Lechler, R. I., DePaoli, A. M., and O'Rahilly, S. (2002) Beneficial effects of leptin on obesity, T cell hyporesponsiveness, and neuroendocrine/metabolic dysfunction of human congenital leptin deficiency. *J. Clin. Invest.* **110,** 1093–1103 CrossRef Medline
15. Farooqi, I. S., and O'Rahilly, S. (2014) 20 years of leptin: human disorders of leptin action. *J. Endocrinol.* **223,** T63–T70 CrossRef Medline
16. Funcke, J. B., von Schnurbein, J., Lennerz, B., Lahr, G., Debatin, K. M., Fischer-Posovszky, P., and Wabitsch, M. (2014) Monogenic forms of childhood obesity due to mutations in the leptin gene. *Mol. Cell Pediatr.* **1,** 3 CrossRef Medline
17. Wabitsch, M., Pridzun, L., Ranke, M., von Schnurbein, J., Moss, A., Brandt, S., Kohlsdorf, K., Moepps, B., Schaab, M., Funcke, J. B., Gierschik, P., Fischer-Posovszky, P., Flehmig, B., and Kratzsch, J. (2017) Measurement of immunofunctional leptin to detect and monitor patients with functional leptin deficiency. *Eur. J. Endocrinol.* **176,** 315–322 CrossRef Medline
18. Niv-Spector, L., Shpilman, M., Grupi, A., and Gertler, A. (2010) The obese phenotype-inducing N82K mutation in human leptin disrupts receptor-binding and biological activity. *Mol. Genet. Metab.* **100,** 193–197 CrossRef Medline
19. Wabitsch, M., Funcke, J. B., Lennerz, B., Kuhnle-Krahl, U., Lahr, G., Debatin, K. M., Vatter, P., Gierschik, P., Moepps, B., and Fischer-Posovszky, P. (2015) Biologically inactive leptin and early-onset extreme obesity. *N. Engl. J. Med.* **372,** 48–54 CrossRef Medline
20. Myers, M. G., Cowley, M. A., and Münzberg, H. (2008) Mechanisms of leptin action and leptin resistance. *Annu. Rev. Physiol.* **70,** 537–556 CrossRef Medline
21. Frühbeck, G. (2001) A heliocentric view of leptin. *Proc. Nutr. Soc.* **60,** 301–318 CrossRef Medline
22. Baratta, M. (2002) Leptin: from a signal of adiposity to a hormonal mediator in peripheral tissues. *Med. Sci. Monit.* **8,** RA282–RA292 Medline
23. Frühbeck, G. (2002) Peripheral actions of leptin and its involvement in disease. *Nutr. Rev.* **60,** S47–S55; discussion S68–S84, 85–87 CrossRef Medline
24. Muoio, D. M., and Lynis Dohm, G. (2002) Peripheral metabolic actions of leptin. *Best Pract. Res. Clin. Endocrinol. Metab.* **16,** 653–666 CrossRef Medline

25. Bjørbaek, C., and Kahn, B. B. (2004) Leptin signaling in the central nervous system and the periphery. *Recent Prog. Horm. Res.* **59,** 305–331 CrossRef Medline
26. Mancour, L. V., Daghestani, H. N., Dutta, S., Westfield, G. H., Schilling, J., Oleskie, A. N., Herbstman, J. F., Chou, S. Z., and Skiniotis, G. (2012) Ligand-induced architecture of the leptin receptor signaling complex. *Mol. Cell* **48,** 655–661 CrossRef Medline
27. Moharana, K., Zabeau, L., Peelman, F., Ringler, P., Stahlberg, H., Tavernier, J., and Savvides, S. N. (2014) Structural and mechanistic paradigm of leptin receptor activation revealed by complexes with wild-type and antagonist leptins. *Structure* **22,** 866–877 CrossRef Medline
28. Peelman, F., Zabeau, L., Moharana, K., Savvides, S. N., and Tavernier, J. (2014) 20 years of leptin: insights into signaling assemblies of the leptin receptor. *J. Endocrinol.* **223,** T9–T23 CrossRef Medline
29. Fong, T. M., Huang, R. R., Tota, M. R., Mao, C., Smith, T., Varnerin, J., Karpitskiy, V. V., Krause, J. E., and Van der Ploeg, L. H. (1998) Localization of leptin binding domain in the leptin receptor. *Mol. Pharmacol.* **53,** 234–240 CrossRef Medline
30. Zabeau, L., Defeau, D., Van der Heyden, J., Iserentant, H., Vandekerckhove, J., and Tavernier, J. (2004) Functional analysis of leptin receptor activation using a Janus kinase/signal transducer and activator of transcription complementation assay. *Mol. Endocrinol.* **18,** 150–161 CrossRef Medline
31. Yang, R., and Barouch, L. A. (2007) Leptin signaling and obesity: cardiovascular consequences. *Circ. Res.* **101,** 545–559 CrossRef Medline
32. Carpenter, B., Hemsworth, G. R., Wu, Z., Maamra, M., Strasburger, C. J., Ross, R. J., and Artymiuk, P. J. (2012) Structure of the human obesity receptor leptin-binding domain reveals the mechanism of leptin antagonism by a monoclonal antibody. *Structure* **20,** 487–497 CrossRef Medline
33. Haglund, E., Sułkowska, J. I., He, Z., Feng, G. S., Jennings, P. A., and Onuchic, J. N. (2012) The unique cysteine knot regulates the pleotropic hormone leptin. *PLoS One* **7,** e45654 CrossRef Medline
34. Connolly, M. L., Kuntz, I. D., and Crippen, G. M. (1980) Linked and threaded loops in proteins. *Biopolymers* **19,** 1167–1182 CrossRef Medline
35. Haglund, E., Sulkowska, J. I., Noel, J. K., Lammert, H., Onuchic, J. N., and Jennings, P. A. (2014) Pierced lasso bundles are a new class of knot-like motifs. *PLoS Comput. Biol.* **10,** e1003613 CrossRef Medline
36. Haglund, E. (2015) Engineering covalent loops in proteins can serve as an on/off switch to regulate threaded topologies. *J. Phys. Condens. Matter* **27,** 354107 CrossRef Medline
37. Dabrowski-Tumanski, P., Niemyska, W., Pasznik, P., and Sulkowska, J. I. (2016) LassoProt: server to analyze biopolymers with lassos. *Nucleic Acids Res.* **44,** W383–W389 CrossRef Medline
38. Niemyska, W., Dabrowski-Tumanski, P., Kadlof, M., Haglund, E., Sulkowski, P., and Sułkowska, J. I. (2016) Complex lasso: new entangled motifs in proteins. *Sci. Rep.* **6,** 36895 CrossRef Medline
39. Haglund, E., Pilko, A., Wollman, R., Jennings, P. A., and Onuchic, J. N. (2017) Pierced lasso topology controls function in leptin. *J. Phys. Chem. B* **121,** 706–718 CrossRef Medline
40. Pigeyre, M., Yazdi, F. T., Kaur, Y., and Meyre, D. (2016) Recent progress in genetics, epigenetics and metagenomics unveils the pathophysiology of human obesity. *Clin. Sci.* (*Lond.*) **130,** 943–986 CrossRef Medline
41. Zhang, Y., Proenca, R., Maffei, M., Barone, M., Leopold, L., and Friedman, J. M. (1994) Positional cloning of the mouse obese gene and its human homologue. *Nature* **372,** 425–432 CrossRef Medline
42. Montague, C. T., Farooqi, I. S., Whitehead, J. P., Soos, M. A., Rau, H., Wareham, N. J., Sewter, C. P., Digby, J. E., Mohammed, S. N., Hurst, J. A., Cheetham, C. H., Earley, A. R., Barnett, A. H., Prins, J. B., and O'Rahilly, S. (1997) Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature* **387,** 903–908 CrossRef Medline
43. Moon, B. C., and Friedman, J. M. (1997) The molecular basis of the obese mutation in ob2J mice. *Genomics* **42,** 152–156 CrossRef Medline
44. Boute, N., Zilberfarb, V., Camoin, L., Bonnafous, S., Le Marchand-Brustel, Y., and Issad, T. (2004) The formation of an intrachain disulfide bond in the leptin protein is necessary for efficient leptin secretion. *Biochimie* (*Paris*) **86,** 351–356 CrossRef Medline
45. Chekhranova, M. K., Karpova, S. K., Iatsyshina, S. B., and Pankov Iu, A. (2008) [A new mutation c.422C>G (p.S141C) in homo- and heterozygous forms of the human leptin gene]. *Bioorg. Khim.* **34,** 854–856 Medline
46. Mazen, I., El-Gammal, M., Abdel-Hamid, M., and Amr, K. (2009) A novel homozygous missense mutation of the leptin gene (N103K) in an obese Egyptian patient. *Mol. Genet. Metab.* **97,** 305–308 CrossRef Medline
47. Hong, C. J., Tsai, P. J., Cheng, C. Y., Chou, C. K., Jheng, H. F., Chuang, Y. C., Yang, C. N., Lin, Y. T., Hsu, C. W., Cheng, I. H., Chen, S. Y., Tsai, S. J., Liou, Y. J., and Tsai, Y. S. (2010) ENU mutagenesis identifies mice with morbid obesity and severe hyperinsulinemia caused by a novel mutation in leptin. *PLoS One* **5,** e15333 CrossRef Medline
48. Fatima, W., Shahid, A., Imran, M., Manzoor, J., Hasnain, S., Rana, S., and Mahmood, S. (2011) Leptin deficiency and leptin gene mutations in obese children from Pakistan. *Int. J. Pediatr. Obes.* **6,** 419–427 CrossRef Medline
49. Saeed, S., Butt, T. A., Anwer, M., Arslan, M., and Froguel, P. (2012) High prevalence of leptin and melanocortin-4 receptor gene mutations in children with severe obesity from Pakistani consanguineous families. *Mol. Genet. Metab.* **106,** 121–126 CrossRef Medline
50. Thakur, S., Kumar, A., Dubey, S., Saxena, R., Peters, A. N., and Singhal, A. (2014) A novel mutation of the leptin gene in an Indian patient. *Clin. Genet.* **86,** 391–393 CrossRef Medline
51. Wabitsch, M., Funcke, J. B., von Schnurbein, J., Denzer, F., Lahr, G., Mazen, I., El-Gammal, M., Denzer, C., Moss, A., Debatin, K. M., Gierschik, P., Mistry, V., Keogh, J. M., Farooqi, I. S., Moepps, B., and Fischer-Posovszky, P. (2015) Severe early-onset obesity due to bioinactive leptin caused by a p.N103K mutation in the leptin gene. *J. Clin. Endocrinol. Metab.* **100,** 3227–3230 CrossRef Medline
52. Altawil, A. S., Mawlawi, H. A., Alghamdi, K. A., and Almijmaj, F. F. (2016) A novel homozygous frameshift mutation in exon 2 of LEP gene associated with severe obesity: a case report. *Clin. Med. Insights Pediatr.* **10,** 115–118 Medline
53. Peelman, F., Van Beneden, K., Zabeau, L., Iserentant, H., Ulrichts, P., Defeau, D., Verhee, A., Catteeuw, D., Elewaut, D., and Tavernier, J. (2004) Mapping of the leptin binding sites and design of a leptin antagonist. *J. Biol. Chem.* **279,** 41038–41046 CrossRef Medline
54. Iserentant, H., Peelman, F., Defeau, D., Vandekerckhove, J., Zabeau, L., and Tavernier, J. (2005) Mapping of the interface between leptin and the leptin receptor CRH2 domain. *J. Cell Sci.* **118,** 2519–2527 CrossRef Medline
55. Fischer-Posovszky, P., von Schnurbein, J., Moepps, B., Lahr, G., Strauss, G., Barth, T. F., Kassubek, J., Mühleder, H., Möller, P., Debatin, K. M., Gierschik, P., and Wabitsch, M. (2010) A new missense mutation in the leptin gene causes mild obesity and hypogonadism without affecting T cell responsiveness. *J. Clin. Endocrinol. Metab.* **95,** 2836–2840 CrossRef Medline
56. Gong, D. W., Bi, S., Pratley, R. E., and Weintraub, B. D. (1996) Genomic structure and promoter analysis of the human obese gene. *J. Biol. Chem.* **271,** 3971–3974 CrossRef Medline
57. Zhang, F., Basinski, M. B., Beals, J. M., Briggs, S. L., Churgay, L. M., Clawson, D. K., DiMarchi, R. D., Furman, T. C., Hale, J. E., Hsiung, H. M., Schoner, B. E., Smith, D. P., Zhang, X. Y., Wery, J. P., and Schevitz, R. W. (1997) Crystal structure of the obese protein leptin-E100. *Nature* **387,** 206–209 CrossRef Medline
58. Murphy, R., and Tsai, A. (eds) (2006) *Misbehaving Proteins: Protein (Mis) Folding, Aggregation, and Stability*, pp. 331–350, Springer, New York
59. Whitford, P. C., Noel, J. K., Gosavi, S., Schug, A., Sanbonmatsu, K. Y., and Onuchic, J. N. (2009) An all-atom structure-based potential for proteins: bridging minimal models with all-atom empirical forcefields. *Proteins* **75,** 430–441 CrossRef Medline
60. Baxter, E. L., Zuris, J. A., Wang, C., Vo, P. L., Axelrod, H. L., Cohen, A. E., Paddock, M. L., Nechushtai, R., Onuchic, J. N., and Jennings, P. A. (2013) Allosteric control in a metalloprotein dramatically alters function. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 948–953 CrossRef Medline
61. Pires, D. E., Ascher, D. B., and Blundell, T. L. (2014) mCSM: predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics* **30,** 335–342 CrossRef Medline

62. Sormanni, P., Aprile, F. A., and Vendruscolo, M. (2015) The CamSol method of rational design of protein mutants with enhanced solubility. *J. Mol. Biol.* **427,** 478–490 CrossRef Medline
63. Blazer, D. G., and Hernandez, L. M. (2006) *Genes, Behavior, and the Social Environment: Moving Beyond the Nature/Nurture Debate*, National Academies Press, Washington, D. C.
64. Verploegen, S. A., Plaetinck, G., Devos, R., Van der Heyden, J., and Guisez, Y. (1997) A human leptin mutant induces weight gain in normal mice. *FEBS Lett.* **405,** 237–240 CrossRef Medline
65. Shabana and Hasnain, S. (2016) The p.N103K mutation of leptin (LEP) gene and severe early onset obesity in Pakistan. *Biol. Res.* **49,** 23 CrossRef Medline
66. Kratzsch, J., Berthold, A., Lammert, A., Reuter, W., Keller, E., and Kiess, W. (2002) A rapid, quantitative immunofunctional assay for measuring human leptin. *Horm. Res.* **57,** 127–132 Medline
67. Khan, S., and Vihinen, M. (2007) Spectrum of disease-causing mutations in protein secondary structures. *BMC Struct. Biol.* **7,** 56 CrossRef Medline
68. Noel, J. K., Levi, M., Raghunathan, M., Lammert, H., Hayes, R. L., Onuchic, J. N., and Whitford, P. C. (2016) SMOG 2: a versatile software package for generating structure-based models. *PLoS Comput. Biol.* **12,** e1004794 CrossRef Medline
69. Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., and Bateman, A. (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44,** D279–D285 CrossRef Medline
70. Gierut, A. M., Niemyska, W., Dabrowski-Tumanski, P., Sulkowski, P., and Sulkowska, J. I. (2017) PyLasso: a PyMOL plugin to identify lassos. *Bioinformatics* **33,** 3819–3821 CrossRef Medline