THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 9, pp. 5845–5851, March 3, 2006

Solution Structure and Novel Insights into the Determinants of the Receptor Specificity of Human Relaxin-3*

Received for publication, October 14, 2005, and in revised form, December 12, 2005 Published, JBC Papers in Press, December 19, 2005, DOI 10.1074/jbc.M511210200

K. Johan Rosengren[‡§], Feng Lin[¶], Ross A. D. Bathgate[¶], Geoffrey W. Tregear[¶], Norelle L. Daly[‡], John D. Wade[¶], and David J. Craik[‡1]

From the [‡]Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland 4072, Australia, the [¶]Howard Florey Institute, University of Melbourne, Victoria 3010, Australia, and the [§]Department of Chemistry and Biomedical Sciences, University of Kalmar, SE-392 81 Kalmar, Sweden

Relaxin-3 is the most recently discovered member of the relaxin family of peptide hormones. In contrast to relaxin-1 and -2, whose main functions are associated with pregnancy, relaxin-3 is involved in neuropeptide signaling in the brain. Here, we report the solution structure of human relaxin-3, the first structure of a relaxin family member to be solved by NMR methods. Overall, relaxin-3 adopts an insulin-like fold, but the structure differs crucially from the crystal structure of human relaxin-2 near the B-chain terminus. In particular, the B-chain C terminus folds back, allowing Trp^{B27} to interact with the hydrophobic core. This interaction partly blocks the conserved RXXXRXXI motif identified as a determinant for the interaction with the relaxin receptor LGR7 and may account for the lower affinity of relaxin-3 relative to relaxin for this receptor. This structural feature is likely important for the activation of its endogenous receptor, GPCR135.

Recent developments have caused considerable excitement in the relaxin field, including the discovery of a new member of the relaxin family, relaxin-3 (1), and the identification of several long sought after relaxin receptors (2). Prior to the discovery of the relaxin-3 gene (*RLN3*), only one relaxin gene had been characterized in most mammals, with the exception of humans and higher primates, in which two separate genes, *RLN1* and *RLN2*, were known (3, 4). The product of the human *RLN2* gene, human relaxin-2 (H2 relaxin), is the functional ortholog of the *RLN1* gene product from non-primate species, which is termed relaxin. The function of the product of the human *RLN1* gene, human relaxin-1 (H1 relaxin), is unknown; and indeed, a native H1 relaxin peptide has not been isolated (5). Hence, throughout this work, "relaxin" will refer to the pregnancy hormones H2 relaxin and non-primate relaxin. Relaxin-3 is the ancestor of the entire relaxin peptide family (5), and *RLN3* genes have been identified in all mammalian genomes as well as in the genomes of chicken, frog, and various fish species. In contrast, the *RLN1* gene is found only in mammals. Fig. 1 shows a sequence comparison of the product of the human *RLN3* gene, human relaxin-3 (H3 relaxin); H2 and H1 relaxins; and relaxin-3 orthologs from other species. Interestingly, the relaxin-3 sequences are highly conserved between species, in contrast to relaxin, which shows considerable sequence variation.

Relaxin has long been regarded as a hormone mainly associated with pregnancy. Produced in the corpus luteum and/or placenta of most mammals, it has numerous pregnancy-specific actions, including the remodeling of the reproductive tract and preparation of the mammary apparatus for lactation (6). However, relaxin also has important physiological roles outside of pregnancy. It inhibits collagen biosynthesis and promotes collagen breakdown (7–9) and causes vasodilatation in various tissues (10). Interestingly, the expression pattern of relaxin-3 differs significantly from that of relaxin, suggesting a distinctly different physiological role. The highest expression of relaxin-3 in all species examined to date, including humans, is in the brain. In rats (11) and mice (1), the expression is localized to a specific region of the dorsal tegmental region of the pons called the nucleus incertus. Anatomical studies suggest that this nucleus is involved in a midbrain behavior control network that influences circuits regulating locomotion, attention, and learning processes and that responds to stress-related neuroendocrine signals (12). Mapping studies have demonstrated that relaxin-3-immunoreactive nerve fibers emanating from the nucleus incertus innervate numerous regions of the brain, suggesting that the nucleus incertus utilizes relaxin-3 as a neurotransmitter (13). These data, coupled with the remarkable conservation of the relaxin-3 peptide from fish to humans, suggest that relaxin-3 has highly important, but presently unknown, central actions. Indeed, a very recent study suggests that relaxin-3 is involved in appetite regulation (14), and another study has demonstrated that relaxin-3 mRNA production is increased in response to water-restraint stress (13), suggesting a role in the stress response.

Members of the relaxin family are structurally similar to insulin, comprising two peptide chains that are linked by two disulfide bonds, with the A-chain containing a third, intrachain disulfide bond. However, unlike insulin, whose receptor is a tyrosine kinase, members of the relaxin peptide family bind to G-protein-coupled receptors (GPCRs).[2] The first receptors identified responding to relaxin stimuli were the leucine-rich repeat-containing GPCRs LGR7 and LGR8 (2), which both have low nanomolar affinity for relaxin and which are capable of mediating the actions of relaxin through a cAMP-dependent pathway (2). Interestingly, relaxin-3 also activates LGR7, albeit with a lower affinity than relaxin, but has a significantly reduced affinity for LGR8 (15). Recent work has identified two additional orphan GPCRs, GPCR135 and GPCR142, which also respond to relaxin-3 activation (16, 17). On the basis of the higher affinity of relaxin-3 for GPCR135 and their coexpression in regions of the brain, it was concluded that relaxin-3 is the endogenous ligand for GPCR135, whereas the primary ligands for LGR8 and GPCR142 are two other members of the relaxin peptide family, INSL3 (insulin-like peptide-3 or Leydig insulin-like peptide) (18) and INSL5 (insulin-like peptide-5) (19), respectively. The cross-reactivity of

* This work was supported in part by National Health and Medical Research Council Project Grants 350284 (to J. D. W.) and 300012 (to R. A. D. B. and J. D. W.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Australian Research Council Professorial Fellow. To whom correspondence should be addressed. Tel.: 61-7-3346-2019; Fax: 61-7-3346-2029; E-mail: d.craik@imb.uq.edu.au.

[2] The abbreviations used are: GPCRs, G-protein-coupled receptors; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; HPLC, high pressure liquid chromatography; NOE, nuclear Overhauser effect.

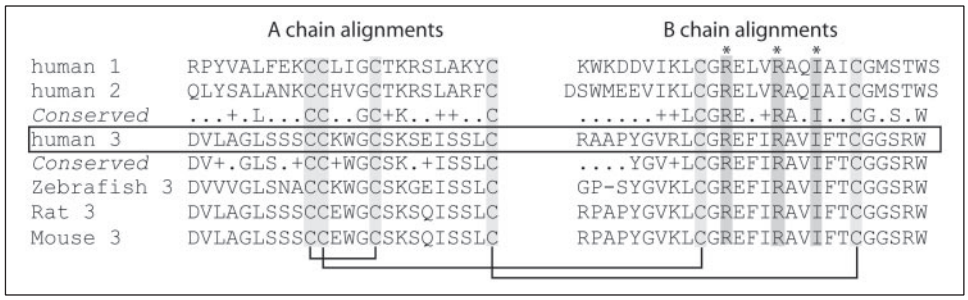

FIGURE 1. **Sequence alignment of H3 relaxin (*boxed*) with H1 and H2 relaxins and relaxin-3 orthologs from other species.** Conserved residues are indicated in single-letter amino acid codes, and conservative changes are marked with *plus signs*. The conserved cysteine residues *shaded*, and their connectivities indicated by *brackets*. The characteristic relaxin binding motif R*XXX*R*XX*(I/V) is *shaded* and marked with *asterisks*.

these hormone-receptor signaling systems has raised questions about which factors are responsible for the selectivity of these ligand-receptor pairings. Moreover, as relaxin-3, GPCR135, and LGR7 are all expressed in the brain, understanding the significance of the interaction of relaxin-3 with LGR7 will be crucial in determining its function.

In this study, we provide a crucial piece of the puzzle by presenting the solution structure of the human form of the most recently discovered member of this hormone family, relaxin-3. The structure reveals a relaxin/insulin fold, but with an unusual conformation of the C-terminal region of the B-chain. The structural data provide insights into the receptor specificities of this new member of the relaxin peptide family.

EXPERIMENTAL PROCEDURES

Peptide Synthesis—The A- and B-chains were assembled as C-terminal peptide acids or amides by Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) solid-phase synthesis on a hydroxymethylphenoxyacetyl- or 5-(4-aminomethyl-3,5-dimethoxyphenoxy)valeryl-polyethylene glycol/polystyrene support using the following selective cysteine *S*-protection: $Cys^{A10/A15}$, trityl; $Cys^{A11/B11}$, acetamidomethyl; Cys^{A24}, *t*-butyl; and Cys^{B22}, trityl. Following conventional trifluoroacetic acid cleavage in the presence of scavengers, the intramolecular disulfide bond of the A-chain was formed by aeration, and Cys^{A24} (*t*-butyl) was displaced with *S*-pyridinyl by reaction with 2,2′-dipyridyl disulfide in trifluoromethanesulfonic acid. Combination of the peptide with the B-chain occurred by thiolysis, after which the third and final disulfide bond between A11 and B10 was formed by iodolysis. The purity of the peptide was confirmed by MALDI-TOF mass spectrometry (for example, peptide acid: theory, 5500.5; and found, 5499.6) and analytical reversed-phase HPLC.

Biological Activity Assays—H3 relaxin acid and amide were tested for their ability to activate LGR7, LGR8, and GPCR135. Stably transfected human embryonic kidney 293T cells expressing human LGR7 or LGR8 (15) and Chinese hamster ovary K1 cells stably expressing human GPCR135 (20) were stimulated for 30 min with various concentrations of H3 relaxin acid or amide. LGR7 and LGR8 were tested in parallel with H2 relaxin and human INSL3, respectively (both at 100 nM), and GPCR135 with forskolin (5 μM) to determine maximum cellular cAMP response. cAMP accumulation was measured in cell lysates as described previously (15).

NMR Spectroscopy—Pulsed-field gradient NMR diffusion experiments were performed with a two-dimensional sequence using stimulated echo longitudinal encode-decode (21). The lengths of all pulses and delays were held constant, and 32 spectra were acquired with the strength of the diffusion gradient varying between 2 and 95% of its maximum value. The lengths of the diffusion gradient and the stimulated echo were optimized for each sample to give a total decay in the protein signal of ~90%. Dioxane was added to the samples to a final concentration of 0.2 mM and used as an internal standard (22). Samples prepared for structure determination contained ~1 mM peptide dissolved in 90% H_2O and 10 or 100% (v/v) D_2O at pH 3.0. Spectra were recorded at 290, 298, and 303 K on a Bruker Avance 600-MHz spectrometer or on a Bruker DMX 750-MHz spectrometer. Two-dimensional experiments recorded included double-quantum filtered COSY; total correlation spectroscopy using an MLEV-17 spin lock sequence with a mixing time of 80 ms; and nuclear Overhauser effect (NOE) correlation spectroscopy with mixing times of 100, 150, and 200 ms. Spectra were generally acquired with 4096 complex data points in F2 and 512 increments in the F1 dimension over a spectral width of 14 ppm. Slowly exchanging NH protons were detected by acquiring a series of one-dimensional and total correlation spectra of the fully protonated peptide immediately after dissolution in D_2O. As most amides disappeared within the first 2 h, resonances still visible after 2 h were considered to be protected from the solvent by hydrogen bonding. Spectra were processed on a Silicon Graphics Octane workstation using XWIN-NMR (Bruker BioSpin Corp.). The F1 dimension was generally zero-filled to 1024 real data points, and 90° phase-shifted sine bell window functions were applied before Fourier transformation. Chemical shifts were referenced to 2,2-dimethyl-2-silapentanesulfonic acid at 0.00 ppm.

Distance restraints were derived primarily from a 100-ms NOE correlation spectrum recorded at 298 K and 600 MHz. Cross-peaks were assigned and integrated in XEASY and converted to distance restraints using CYANA. Distance restraints for cross-peaks that could not be unambiguously assigned were introduced into structure calculations as ambiguous restraints. Backbone dihedral restraints were inferred from $^3J_{NH\text{-}H\alpha}$ coupling constants. Because of the generally broad lines of many amide protons, coupling constants were, in some cases, estimated from a combination of apparent coupling constants, peak intensities, and consistency with preliminary structures. The dihedral angle ϕ was restrained to $-120 \pm 40°$ for $^3J_{NH\text{-}H\alpha} > 8$ Hz and to $-60 \pm 30°$ for $^3J_{NH\text{-}H\alpha} < 5$ Hz. Additional ϕ angle restraints of $-100 \pm 80°$ were included where a positive angle could be excluded based on a strong sequential $H\alpha_{i-1}$-HN_i NOE compared with the intraresidue $H\alpha_i$-HN_i NOE. Side chain χ_1 angles and stereospecific assignments were determined on the basis of observed NOE and $^3J_{H\alpha\text{-}H\beta}$ coupling patterns. For a predicted t^2g^3 side chain conformation, the χ_1 angles were restrained to $-60 \pm 30°$ (residues Trp^{A13}, Glu^{A19}, Cys^{A24}, Arg^{B8}, Leu^{B9}, Cys^{B10}, and Cys^{B22}); and for a g^2t^3 conformation, the angles were constrained to $180 \pm 30°$ (residues Lys^{A12}, Lys^{A17}, Leu^{A23}, Phe^{B14}, and Arg^{B26}). No residues could be confirmed to be in the g^2g^3 conformation based on experimental data. Additional stereospecific assignments of the methyl groups for Val^{B18} and Leu^{A23} were determined based on their NOE patterns. Hydrogen bonds were included in the structure calculations for all amide protons concluded to be slowly exchanging only after a suitable acceptor could be identified in the preliminary structures. In all cases, these hydrogen bonds were found between the backbone atoms within the elements of secondary structure. Three-dimensional structures were calculated using simulated annealing and energy minimization protocols from ARIA (23) in the program CNS (24). The protocol

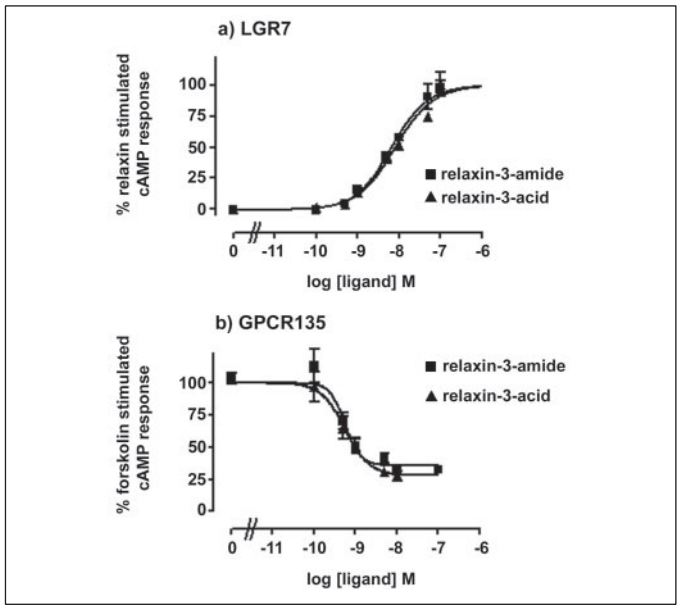

FIGURE 2. **cAMP accumulation in response to H3 relaxin amide and acid in stably transfected LGR7-expressing (*a*) and GPCR135-expressing (*b*) cells.** Activities are plotted as the percent maximum H2 relaxin- and forskolin-stimulated responses, respectively. Stimulation of LGR7 resulted in increases in cellular cAMP, whereas stimulation of GPCR135 resulted in decreases in cellular cAMP. The activities of the acid and amide forms of the peptide were identical for both receptors.

involves a high temperature phase with 4000 steps of 0.015 ps of torsion angle dynamics; a cooling phase with 4000 steps of 0.015 ps of torsion angle dynamics during which the temperature is lowered to 0 K; and finally, an energy minimization phase with 5000 steps of Powell minimization. The refinement in explicit water involves the following steps: heating to 500 K via steps of 100 K, each with 50 steps of 0.005 ps of Cartesian dynamics; 2500 steps of 0.005 ps of Cartesian dynamics at 500 K; and a cooling phase in which the temperature is lowered in steps of 100 K, each with 2500 steps of 0.005 ps of Cartesian dynamics. Finally, the structures were minimized with 2000 steps of Powell minimization.

Protein structures were analyzed using PROMOTIF and PROCHECK and displayed using MOLMOL. Ramachandran analysis showed that 83% of the residues are in the most favored regions, with the remaining in the additionally allowed (15%) and generously allowed (2%). The coordinates representing the solution structure of H3 relaxin and the experimental restraints have been submitted to the Protein Data Bank (Code 2FHW).

RESULTS

Peptide Synthesis—H3 relaxin was prepared as both its C-terminal acid and amide forms by solid-phase peptide synthesis of the separate A- and B-chains together with regioselective disulfide bond formation (25). Chemical characterization by reversed-phase HPLC and MALDI-TOF mass spectrometry confirmed the homogeneity of the products. Tryptic mapping followed by MALDI-TOF mass spectrometry showed that the disulfide bonds were in the correct insulin-like pairings.

Biological Activity—H3 relaxin acid was tested for its ability to stimulate human LGR7, LGR8, and GPCR135 in parallel with the H3 relaxin amide peptide. As shown in Fig. 2, the activities of the peptides were identical. In LGR7-expressing cells, both peptides stimulated cAMP accumulation to levels similar to those achieved with H2 relaxin (H3 relaxin acid $pEC_{50} = 8.08 \pm 0.07$ ($n = 4$) and H3 relaxin amide $pEC_{50} = 8.19 \pm 0.078$ ($n = 3$)). In GPCR135-expressing cells, both peptides decreased forskolin-stimulated cAMP accumulation (H3 relaxin acid $pEC_{50} = 9.30 \pm 0.082$ ($n = 3$) and H3 relaxin amide $pEC_{50} = 9.25 \pm 0.092$ ($n = 3$)). Both peptides were unable to stimulate cAMP accumulation in LGR8-expressing cells at concentrations up to 1 μM (data not shown).

NMR Diffusion Measurements—In light of the tendency of members of the insulin superfamily to aggregate, including H2 relaxin and insulin, it was important to establish whether any multimers of H3 relaxin were present under the conditions used for the NMR studies. This was assessed by measuring translational diffusion using pulsed-field gradient experiments with dioxane (hydrodynamic radius of 2.12 Å) as an internal standard. No significant difference in the hydrodynamic radius of H3 relaxin was observed between concentrations of 0.1 and 1 mM (14.2 and 14.4 Å, respectively), indicating that there was no concentration-dependent aggregation. On the basis of the equation derived by Wilkins *et al.* (22) ($R_h = (4.75 \pm 1.11)N^{0.29 \pm 0.02}$, where N is the number of residues in the protein and R_h is the hydrodynamic radius in Angstroms), the expected values for a H3 relaxin monomer (51 residues) and a H3 relaxin dimer (102 residues) would be ~14.8 and ~18.2 Å, respectively, confirming that H3 relaxin is monomeric under the conditions used for these studies.

NMR Assignments and Structure Determination of H3 Relaxin—For the structural analysis of H3 relaxin, extensive two-dimensional NMR spectral data were recorded at 600 and 750 MHz on a sample containing ~1 mM peptide. The spectral data were of high quality with excellent signal dispersion indicative of a well structured peptide. Resonance assignments were achieved using two-dimensional sequential assignment strategies, which, after analysis of spectra of both the H3 relaxin acid and amide forms at several temperatures, resulted in complete assignments for the peptide backbone and nearly complete assignment for side chain resonances. Interestingly, a number of residues showed broad lines; and in particular, significant line broadening was observed for a set of resonances, including Ser^{A7}, Cys^{A10}, Cys^{A11}, Arg^{A12}, Trp^{A13}, Gly^{A14}, Cys^{A15}, Cys^{B10}, and Phe^{B14}, suggesting that conformational exchange is present in this region of the molecule. This exchange is likely a result of a pro-*R*/pro-*S* reorientation of one of the Cys^{A10}–Cys^{A15} or Cys^{A11}–Cys^{B10} disulfide bonds. In addition, several weak spin systems arising from a minor conformation were identified. These resonances were found to correspond to the N-terminal region of the B-chain, *i.e.* Ala^{B2}, Ala^{B3}, Pro^{B4}, and Tyr^{B5}, and analysis of NOEs revealed that they were the result of a *cis*/*trans*-isomerization of Pro^{B4}. The *trans*-conformation is the major isomer (~90%), with the *cis*-conformation corresponding to a minor conformation (~10%), as evident from strong sequential $H\alpha_{i-1}$-$H\delta_i$ and $H\alpha_{i-1}$-$H\alpha_i$ NOEs, respectively.

The structural restraints derived from the NMR data and used for structure calculations were based exclusively on data recorded for the native acid form and included interproton distances, backbone and side chain dihedral angles derived from coupling constants, and restraints for hydrogen bonds deduced from amide exchange experiments. Structures were calculated by simulated annealing followed by refinement and energy minimization in explicit solvent (H_2O). Distance restraints were derived from intensities of NOE cross-peaks in an NOE correlation spectrum recorded at 298 K with a mixing time of 100 ms. As expected for a 51-amino acid peptide, a number of cross-peaks were ambiguous at the first stage of the assignment due to degeneracy of chemical shifts, in particular in the methyl region. These ambiguities were resolved using an iterative approach in which preliminary structures were used to guide the assignments and by inclusion of ambiguous restraints. Hydrogen bonds were inferred from amide exchange behavior and introduced into the structure calculations once the preliminary structures indicated a suitable acceptor, which in all cases were found to be within elements of secondary structure.

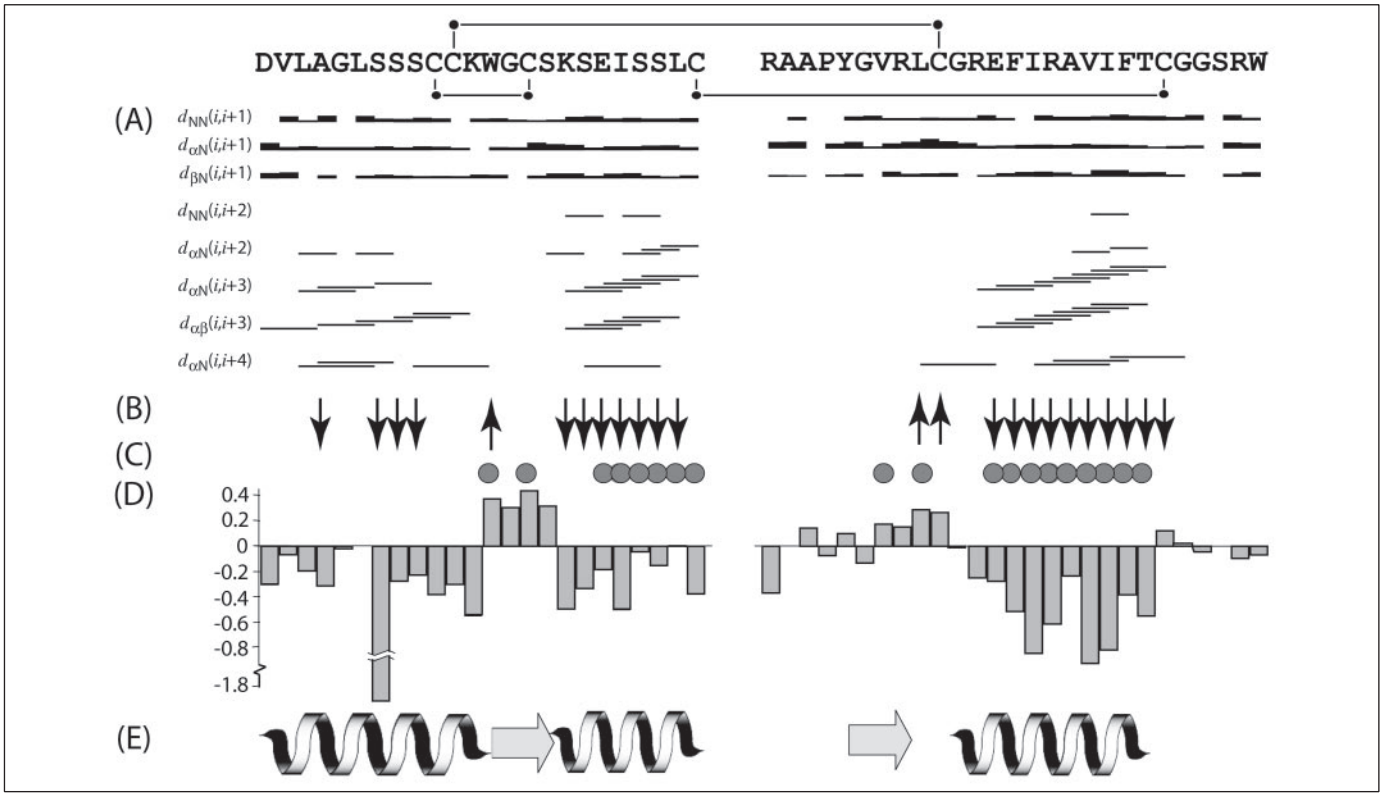

FIGURE 3. **Summary of NMR data on H3 relaxin.** *A*, sequential and medium-range NOEs observed in an NOE correlation spectrum recorded at 600 MHz with a mixing time of 100 ms. The thickness of the *bars* corresponds to the observed NOE intensities. *B*, coupling constants derived from double-quantum filtered COSY spectra. *Downward pointing arrows* correspond to coupling constants <5 Hz, and *upward pointing arrows* correspond to coupling constants >8 Hz. *C*, amide exchange. *Circles* indicate amide protons still visible 1 h after dissolution in 100% D_2O. *D*, secondary Hα shifts in ppm, *i.e.* observed Hα shifts minus random-coil shifts. Stretches of three or more positive or negative values are indicative of extended or helical structure, respectively. *E*, elements of secondary structure in H3 relaxin as inferred from the above data and subsequent structure calculations.

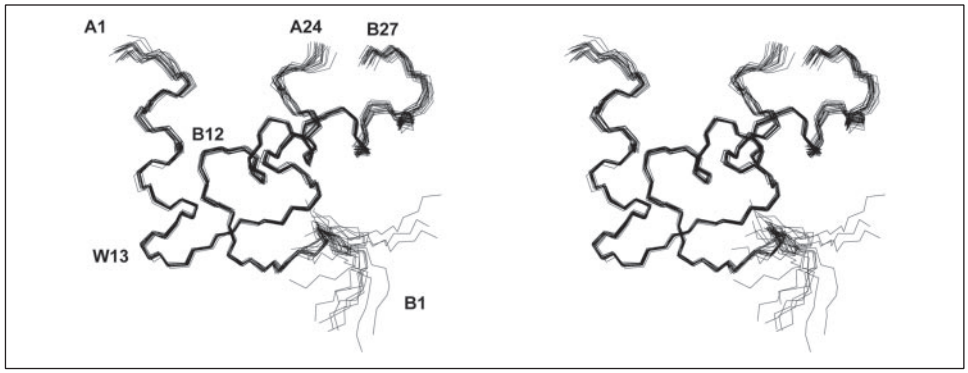

FIGURE 4. **Stereo view of superimposition of the family of 20 structures representing the solution structure of H3 relaxin.** Selected residues are labeled with chain identifiers and residue numbers.

Description of the Three-dimensional Structure of H3 Relaxin—A summary of the NMR data, including sequential and medium-range NOEs, coupling constants, hydrogen exchange, and secondary Hα shifts, is presented in Fig. 3. These data provide a good indication of the presence of secondary structure, as both helical and extended conformations have typical "fingerprints." Helices are generally characterized by a large number of medium-range NOE contacts, small $^3J_{H\alpha\text{-}HN}$ coupling constants, and negative secondary shifts, whereas β-sheets display large $^3J_{H\alpha\text{-}HN}$ coupling constants and positive secondary shifts. The exchange rates between the backbone amide protons and the solvent can be measured by recording the decay in signal after dissolution of the peptide in D_2O and provide information about the presence of hydrogen bonds, as amide protons involved in hydrogen bonds are protected from the solvent and display a slow exchange behavior. Although most amides in H3 relaxin exchange in minutes, a number are still visible >2 h after dissolution in D_2O, and these could in all cases be correlated to hydrogen bonds in the structure.

Fig. 4 shows a superposition of a family of 20 low energy structures representing the solution structure of H3 relaxin. Overall, the structure is well defined, with the exception of the N terminus of the B-chain, which is disordered and likely flexible in solution. The energetic and structural statistics for the structural family are summarized in Table 1. All structures have good energies and covalent geometry as evident from small deviations from ideal bond lengths and angles and the Ramachandran analysis, which showed that 83% of all non-Gly/Pro residues are in the most favored regions and that the remaining residues, all of which are in the more flexible regions of the peptide, are in the additionally allowed regions. It is clear that H3 relaxin adopts a typical relaxin/insulin fold, with the A-chain containing two α-helices (residues A1–A13 and A17–A24) separated by a short β-strand (residues A15–

TABLE 1
NMR and refinement statistics for protein structures

	H3 relaxin
NMR distance and dihedral constraints	
Distance constraints	
Total inter-residue NOEs	766
Sequential ($\|i-j\|=1$)	311
Medium-range ($\|i-j\|<4$)	213
Long-range ($\|i-j\|>5$)	242
Hydrogen bonds	21
Total dihedral angle restraints	
ϕ	33
χ_1	16
Structure statistics	
Violations (mean ± S.D.)	
Distance constraints (>0.2 Å)	0.4 ± 0.60
Dihedral angle constraints (>2°)	0
Maximum distance constraint violation (Å)	0.24
Deviations from idealized geometry	
Bond lengths (Å)	0.00414 ± 0.00011
Bond angles	0.512 ± 0.017°
Impropers	0.412 ± 0.026°
Average pairwise r.m.s.d.[a] (Å)	
Heavy	1.11
Backbone	0.50

[a] Pairwise root mean square deviation (r.m.s.d.) was calculated for 20 refined structures over residues A1–A24 and B3–B27.

A17) and the B-chain containing a second β-strand (residues B5–B7) and an α-helix (residues B12–B22). The A-chain helices lie parallel to each other, forming a U-shaped arrangement. The B-chain helix is placed across the face of the U, roughly perpendicular to the axes of the A-chain helices. Enclosed between them is a significant hydrophobic core involving the side chains of Leu^{A3}, Leu^{A6}, Ile^{A20}, Leu^{A23}, Cys^{A10}, Cys^{A15}, Leu^{B9}, Cys^{B10}, Phe^{B14}, Ile^{B15}, Val^{B18}, Ile^{B19}, and Trp^{B27}. The observation that the Trp^{B27} side chain interacts with the hydrophobic core as evident from a large number of NOEs to Cys^{A24}, Cys^{B22}, Ile^{B15}, Val^{B18}, and Ile^{B19} was surprising because, in the crystal structure of H2 relaxin, this region is disordered, and the Trp extends away from the core, being fully exposed to the solvent. The interaction appears to have become possible because of a shortening of the B-chain helix and a reversal of the peptide backbone due to turns formed by the B23–B26 (GGSR) segment.

DISCUSSION

Solution Structure of H3 Relaxin—In this study, we have determined the three-dimensional solution structure of H3 relaxin. H3 relaxin is well behaved in solution and adopts an insulin-like fold that is braced by the three disulfide bonds conserved throughout the family and that is characterized by three helical segments and a short double-stranded β-sheet. Although the structure is well defined overall, there is evidence for dynamic processes in several regions of the peptide. Most pronounced is the disorder at the N terminus of the B-chain, which is likely due to flexibility in this region. Supporting this suggestion is the observation of *cis*/*trans*-isomerization for Pro^{B4}, which results in two sets of resonances for the amino acids in this region. A similar isomerization has been observed in insulin (26), but there are no data suggesting that this feature is implicated in the biological function. More interesting is the observation of specific broadening of peaks from residues near the Cys^{A10}–Cys^{A15} and Cys^{A11}–Cys^{B10} disulfide bonds, with the extreme example being Cys^{A10}, which is broadened close to the limit of detection, suggesting conformational exchange on a microsecond time scale in this region of the molecule. The most likely explanation for this is a reorientation of the Cys^{A10}–Cys^{A15} disulfide bond conformation as a result of a χ_1 angle flip of one of the cysteines. Such a process has been observed previously in disulfide-rich peptides, including, for example, bovine pancreatic trypsin inhibitor (27, 28). In the H3 relaxin structural family, equal populations of −60 and −180° χ_1 conformations of the Cys^{A15} side chain are observed, supporting the proposal that this disulfide can adopt more than one conformation. Although it is difficult to correlate the two conformations with other differences between the two subpopulations of structures, the proximity of the Cys^{A10}–Cys^{A15} disulfide bond to the side chain of Phe^{B14} means that a reorientation of the disulfide could cause a slight shift in the positioning of the aromatic ring. This could result in turn in a large difference in the chemical shift of neighboring protons between the two conformations due to the ring current effects, which would explain the severe broadening observed for the resonances around Phe^{B14}, including the Cys^{A10}–Cys^{A15} disulfide bond itself.

Amide exchange rates can provide information about the rigidity of a structure in that amide protons involved in strong hydrogen bonds are protected from exchange with the solvent. Among the amides found to be slowly exchanging in H3 relaxin are all amides expected to be involved in hydrogen bonds in the C-terminal α-helix of the A-chain, the α-helix in the B-chain, and between the strands in the β-sheet. However, no slowly exchange amides were detected within the N-terminal α-helix in the A-chain despite a large number of medium-range NOEs, confirming the presence of this α-helix. On the basis of the fast exchange rates, it appears that this α-helix is more dynamic than the other elements of secondary structure in H3 relaxin.

Comparison with the Structures of H2 Relaxin and Insulin—Structural studies of members of the insulin superfamily in solution have been limited by the fact that both H2 relaxin and insulin form dimers or higher aggregates in aqueous solution. High resolution crystal structures of the dimer of H2 relaxin (29) and zinc ion-stabilized insulin hexamers (30–32) have been determined and provide a detailed pictured of the molecular core and elements of secondary structure. However, the biologically active forms are the monomers, and the previous structures do not provide a complete picture, as the dimer interface in both cases involves the receptor-binding face, and it is unclear how the formation of the dimer and crystal packing interactions may affect the active site. In particular, there are questions about the conformation of the C-terminal region of the B-chain, which in insulin forms a β-strand that hydrogen bonds in an intermolecular β-sheet, whereas in H2 relaxin, this region appears to be disordered. To obtain structural data on the insulin monomer and to investigate the determinants of the aggregation, several mutants have been studied, most involving alteration of the C terminus of the B-chain to remove hydrophobic interactions in the dimer (33, 34).

Fig. 5 shows a comparison of the H3 relaxin NMR structure and the x-ray structures of H2 relaxin and insulin, from which it is clear that the overall fold with the characteristic helical segments (residues A1–A13, A17–A24, and B12–B22) is highly conserved and that the main differences are around the termini of the B-chain. The C-terminal tails of the relaxins are shorter, comprising only five residues in H3 relaxin and six residues in H2 relaxin compared with 11 residues in insulin. Although this prevents the formation of a longer extended segment that can interact both with the B-chain helix and with an additional monomer as in insulin, we have shown here that the tail can still fold back to allow interactions between Trp^{B27} and the hydrophobic core, mainly Ile^{B19} and Val^{B18} in H3 relaxin. In contrast, in H2 relaxin, the helix is one turn longer, including three additional helical hydrogen bonds, forcing the tail to extend away from the molecular core. It is interesting to speculate which differences in the primary sequence may be determinants of the conformation of the C-terminal region. The tail is one residue longer in H2 relaxin; but based on the structure of H3 relaxin, an additional res-

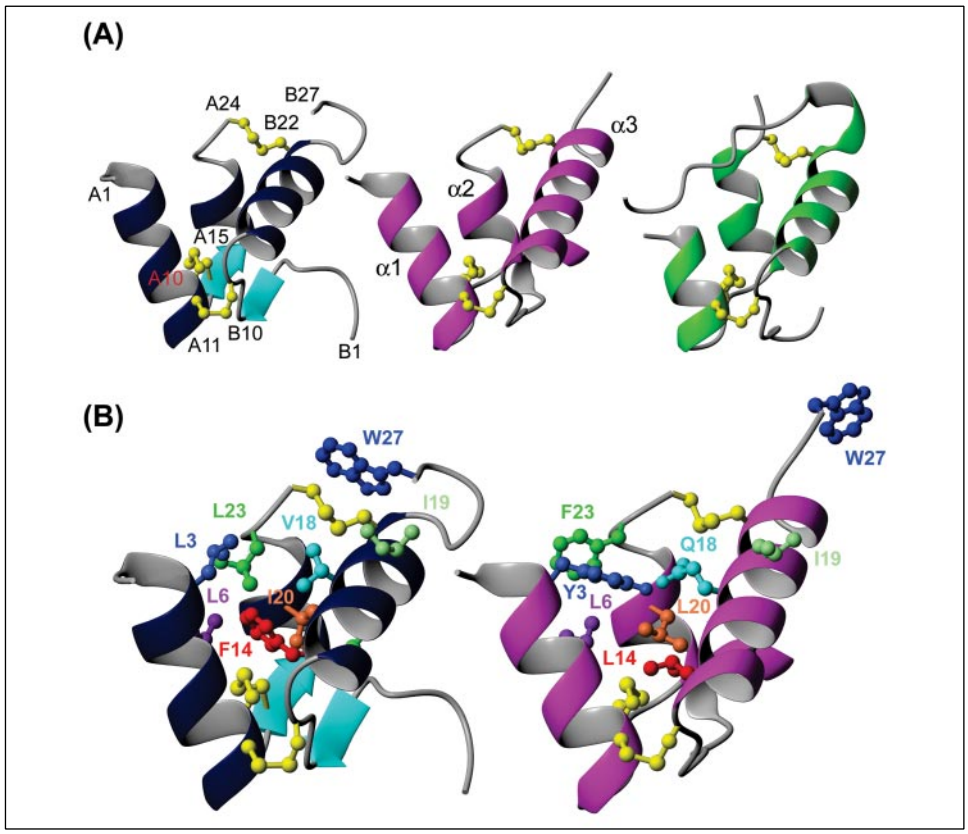

FIGURE 5. *A*, comparison of H3 relaxin (*left*), H2 relaxin (*middle*; Protein Data Bank code 6RLX) (29), and pig insulin (*right*; Protein Data Bank code 4INS) (40). Terminal residues and cysteines are labeled with chain identifiers and residue numbers. A-chain N- and C-terminal helices and the B-chain helix are labeled $\alpha 1$, $\alpha 2$, and $\alpha 3$, respectively. *B*, detailed comparison of the hydrophobic core of H3 relaxin (*left*) and H2 relaxin (*right*). The fold is stabilized by favorable interactions between the side chains of Leu/Tyr^{A3} (*blue*), Leu^{A6} (*purple*), Cys^{A10} (*yellow*), Cys^{A15} (*yellow*), Ile/Leu^{A20} (*orange*), and Leu/Phe^{A23} (*green*) of the A-chains and Phe/Leu^{B14} (*red*), Ala^{B17} (*green*), and Val/Gln^{B18} (*cyan*) of the B-chains of H3/H2 relaxin. In addition, in H3 relaxin, the association of Trp^{B27} with the core extends the core to involve additional interactions with the side chains of Ile^{B19} (*light green*), Cys^{B22} (*yellow*), and Cys^{A24} (*yellow*).

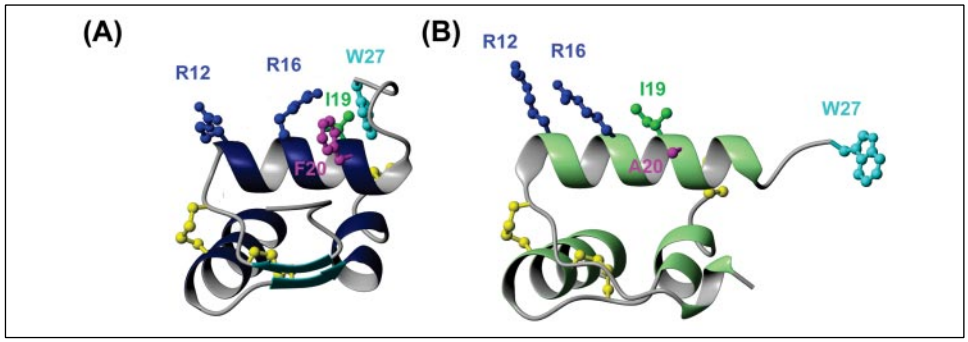

FIGURE 6. **Comparison of the active sites of H3 relaxin (*A*) and H2 relaxin (*B*).** Both peptides have the *RXXXRXX*(I/V) motif shown to be crucial for the interaction with LGR7; however, the C-terminal conformation partly blocks the Ile side chain in H3 relaxin, which likely accounts for the lower affinity for LGR7. A striking feature of H3 relaxin is the highly exposed Phe^{B20} (*purple*) on the same face of the B-chain helix. Being conserved throughout the relaxin-3 family, it is likely that this Phe is involved in the interaction with the relaxin-3 receptor and important for the receptor specificity of relaxin-3.

idue would not appear to interfere with the fold and is unlikely to prevent such a conformation. The main interaction between the tail and the core in H3 relaxin involves hydrophobic contacts between the core and the Trp^{B27} side chain; and as this Trp is conserved, similar interactions would be possible in H2 relaxin. One possibility is that these hydrophobic interactions would be destabilized by the substitution of Val^{B18} in H3 relaxin with Gln^{B18}, which has a more polar side chain. Furthermore, in H2 relaxin, the conformation in which the Trp side chain extends away from the core is stabilized by hydrophobic interactions in a cluster comprising the side chains of Trp^{B27}, Met^{B24}, and Met^{A2}. Neither of these Met residues is conserved in H3 relaxin, which contains Ser and Gly in the equivalent positions; and hence, such an interaction would not be possible. The presence of Met^{B24} in H2 relaxin would also be expected to have a stabilizing effect on the extended C-terminal helix, in contrast to Gly^{B24} in H3 relaxin, which, together with its neighboring residues, Gly^{B23} and Ser^{B25}, would be expected to be strongly helix-breaking.

Implications for Receptor Binding—The observation that, although adopting a well defined structure, H3 relaxin is a rather dynamic molecule is interesting, as studies on insulin have indicated that a significant degree of flexibility is a requirement for binding to the insulin receptor. In particular, an insulin mutant in which the fold was stabilized by the addition of a peptide linker between the N terminus of the A-chain and the C terminus of the B-chain was found to be biologically inactive, despite adopting a native fold (35). In contrast, the des-Phe^{B25} and PT insulin mutants have also been studied by NMR and shown to have significantly destabilized structures (34, 36). These conclusions were based on increased amide exchange rates (34); and in the extreme case of PT insulin, the co-solvent trifluoroethanol, known to stabilize helical structures, had to be added to achieve a well defined structure in solution (36). Strikingly, both these molecules have increased biological activity relative to native insulin (33, 37). Based on these observations, a model for insulin binding to its receptor that involves both detachment of the C-terminal part of the B-chain from the rest of the molecule and unfolding of the N-terminal α-helix has been proposed. Here, we have shown that, in H3 relaxin, the amide exchange rates are also fast, in particular in the N-terminal helix of the A-chain; and in addition, there is evidence for line broadening as a result of conformational exchange around this region. Although the relaxin and relaxin-3 receptors LGR7

and GPCR135, respectively, are GPCRs and the insulin receptor is a tyrosine kinase receptor, and the interactions are thus likely different, it is possible that flexibility may be needed for structural rearrangement in a similar fashion to bind the receptor.

An essential requirement for binding to the relaxin receptor is a conserved R*XXX*R*XX*(I/V) motif that is presented on one face of the B-chain helix (38). Both relaxin and relaxin-3 have this motif and have the ability to stimulate the relaxin receptor LGR7, but relaxin-3 has a 50–100-fold lower activity (1, 15). Recent mutational studies on the relaxin receptor LGR7 revealed a binding surface in the pleated sheet region of the ectodomain comprising two crucial Asp-Glu pairs in the right positions for chelating each of the relaxin active-site Arg residues and, in addition, a cluster of hydrophobic residues equally important for relaxin binding assumed to be clustering with the active-site Ile/Val (38). Interestingly, the structure of H3 relaxin reveals that the C terminus of the B-chain interacts with the molecular core, including the important Ile^{B19} of the active site. As shown in Fig. 6, this interaction partly blocks the accessibility of Ile^{B19} and would interfere with the binding to LGR7. This suggests either that a conformational change, like the detachment of the C-terminal tail proposed for insulin, is needed for H3 relaxin to be able to interact with LGR7 or that the interaction may be different, and this can account for the reduced activity of H3 relaxin relative to H2 relaxin in activating LGR7. It is also apparent that Phe^{B20}, which is presented on the same face of the B-chain helix, is highly exposed to the solvent. As this residue is fully conserved among the relaxin-3 family members, it is interesting to speculate that it may play an important role in the interaction of relaxin-3 with its receptor GPCR135. An additional important feature of the disposition of the C terminus of the B chain is the interaction of H3 relaxin with LGR8. H2 relaxin has been demonstrated to be a high affinity ligand of LGR8, but H3 relaxin is inactive at this receptor (15). It has been demonstrated previously that the C terminus of the B chain in INSL3 (the endogenous ligand of LGR8), in particular the sequence GGPRW, is essential for the interaction of INSL3 with LGR8 (39). The tryptophan residue is essential for LGR8 activity, and this residue is present in the identical position in H2 relaxin. This tryptophan and the entire motif are highly conserved in H3 relaxin, and it was a surprise when H3 relaxin was demonstrated to have no activity at LGR8. Hence, the interaction of Trp^{B27} in H3 relaxin with the hydrophobic core of the molecule demonstrated in this study may explain its lack of activity for LGR8.

Finally, an unusual but characteristic feature of the relaxin peptide family is that high sequence variability is seen between closely related species, whereas distant species show high homology (5). Furthermore, the relationship between primary sequence of relaxins and their biological activity is puzzling. A number of mutational studies have provided valuable data on the residues crucial for the biological activity of several members of the relaxin family. However, although a motif found to be responsible for receptor binding in one hormone may be conserved in other family members, it does not always confer the same activity. In this study, we have shown that, despite the family's highly conserved overall fold, there are differences in the tertiary structures of relaxins and that these differences, not necessarily caused by sequence differences in the active site, may provide an explanation for such observations.

Acknowledgments—We thank Tania Ferraro and Sharon Layfield (Howard Florey Institute) for technical assistance.

REFERENCES

1. Bathgate, R. A. D., Samuel, C. S., Burazin, T. C., Layfield, S., Claasz, A. A., Reytomas, I. G., Dawson, N. F., Zhao, C., Bond, C., Summers, R. J., Parry, L. J., Wade, J. D., and Tregear, G. W. (2002) *J. Biol. Chem.* **277,** 1148–1157
2. Hsu, S. Y., Nakabayashi, K., Nishi, S., Kumagai, J., Kudo, M., Sherwood, O. D., and Hsueh, A. J. (2002) *Science* **295,** 671–674
3. Hudson, P., Haley, J., John, M., Cronk, M., Crawford, R., Haralambidis, J., Tregear, G., Shine, J., and Niall, H. (1983) *Nature* **301,** 628–631
4. Hudson, P., John, M., Crawford, R., Haralambidis, J., Scanlon, D., Gorman, J., Tregear, G., Shine, J., and Niall, H. (1984) *EMBO J.* **3,** 2333–2339
5. Wilkinson, T. N., Speed, T. P., Tregear, G. W., and Bathgate, R. A. D. (2005) *BMC Evolutionary Biology* http://www.biomedcentral.com/1471-2148/5/14
6. Sherwood, O. D. (1994) in *The Physiology of Reproduction* (Knobil, E., and Neill, J. D., eds) pp. 861–1009, Raven Press, Ltd., New York
7. Williams, E. J., Benyon, R. C., Trim, N., Hadwin, R., Grove, B. H., Arthur, M. J., Unemori, E. N., and Iredale, J. P. (2001) *Gut* **49,** 577–583
8. Unemori, E. N., and Amento, E. P. (1990) *J. Biol. Chem.* **265,** 10681–10685
9. Garber, S. L., Mirochnik, Y., Brecklin, C. S., Unemori, E. N., Singh, A. K., Slobodskoy, L., Grove, B. H., Arruda, J. A., and Dunea, G. (2001) *Kidney Int.* **59,** 876–882
10. Bani, D. (1997) *Gen. Pharmacol.* **28,** 13–22
11. Burazin, T. C., Bathgate, R. A. D., Macris, M., Layfield, S., Gundlach, A. L., and Tregear, G. W. (2002) *J. Neurochem.* **82,** 1553–1557
12. Goto, M., Swanson, L. W., and Canteras, N. S. (2001) *J. Comp. Neurol.* **438,** 86–122
13. Tanaka, M., Iijima, N., Miyamoto, Y., Fukusumi, S., Itoh, Y., Ozawa, H., and Ibata, Y. (2005) *Eur. J. Neurosci.* **21,** 1659–1670
14. McGowan, B. M., Stanley, S. A., Smith, K. L., White, N. E., Connolly, M. M., Thompson, E. L., Gardiner, J. V., Murphy, K. G., Ghatei, M. A., and Bloom, S. R. (2005) *Endocrinology* **146,** 3295–3300
15. Sudo, S., Kumagai, J., Nishi, S., Layfield, S., Ferraro, T., Bathgate, R. A. D., and Hsueh, A. J. (2003) *J. Biol. Chem.* **278,** 7855–7862
16. Liu, C., Eriste, E., Sutton, S., Chen, J., Roland, B., Kuei, C., Farmer, N., Jornvall, H., Sillard, R., and Lovenberg, T. W. (2003) *J. Biol. Chem.* **278,** 50754–50764
17. Liu, C., Chen, J., Sutton, S., Roland, B., Kuei, C., Farmer, N., Sillard, R., and Lovenberg, T. W. (2003) *J. Biol. Chem.* **278,** 50765–50770
18. Kumagai, J., Hsu, S. Y., Matsumi, H., Roh, J. S., Fu, P., Wade, J. D., Bathgate, R. A. D., and Hsueh, A. J. (2002) *J. Biol. Chem.* **277,** 31283–31286
19. Liu, C., Kuei, C., Sutton, S., Chen, J., Bonaventure, P., Wu, J., Nepomuceno, D., Kamme, F., Tran, D. T., Zhu, J., Wilkinson, T., Bathgate, R., Eriste, E., Sillard, R., and Lovenberg, T. W. (2005) *J. Biol. Chem.* **280,** 292–300
20. van der Westhuizen, E. T., Sexton, P. M., Bathgate, R. A. D., and Summers, R. J. (2005) *Ann. N. Y. Acad. Sci.* **1041,** 332–337
21. Altieri, A. S., Hinton, D. P., and Byrd, R. A. (1995) *J. Am. Chem. Soc.* **117,** 7566–7567
22. Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) *Biochemistry* **38,** 16424–16431
23. Linge, J. P., and Nilges, M. (1999) *J. Biomol. NMR* **13,** 51–59
24. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
25. Lin, F., Otvos, L., Jr., Kumagai, J., Tregear, G. W., Bathgate, R. A. D., and Wade, J. D. (2004) *J. Pept. Sci.* **10,** 257–264
26. Higgins, K. A., Craik, D. J., Hall, J. G., and Andrews, P. R. (1988) *Drug Des. Delivery* **3,** 159–170
27. Otting, G., Liepinsh, E., and Wuthrich, K. (1993) *Biochemistry* **32,** 3571–3582
28. Grey, M. J., Wang, C., and Palmer, A. G., III (2003) *J. Am. Chem. Soc.* **125,** 14324–14335
29. Eigenbrot, C., Randal, M., Quan, C., Burnier, J., O'Connell, L., Rinderknecht, E., and Kossiakoff, A. A. (1991) *J. Mol. Biol.* **221,** 15–21
30. Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., Mercola, D. A., and Vijayan, M. (1971) *Nature* **231,** 506–511
31. Smith, G. D., Swenson, D. C., Dodson, E. J., Dodson, G. G., and Reynolds, C. D. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 7093–7097
32. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Reynolds, C. D., Smith, G. D., Sparks, C., and Swenson, D. (1989) *Nature* **338,** 594–596
33. Clausen, R., Jorgensen, T. G., Jorgensen, K. H., Johnsen, A. H., Led, J. J., and Josefsen, K. (2002) *Eur. J. Endocrinol.* **147,** 227–233
34. Jorgensen, A. M., Olsen, H. B., Balschmidt, P., and Led, J. J. (1996) *J. Mol. Biol.* **257,** 684–699
35. Derewenda, U., Derewenda, Z., Dodson, E. J., Dodson, G. G., Bing, X., and Markussen, J. (1991) *J. Mol. Biol.* **220,** 425–433
36. Keller, D., Clausen, R., Josefsen, K., and Led, J. J. (2001) *Biochemistry* **40,** 10732–10740
37. Balschmidt, P., and Brange, J. J. V. (1992) *Diabetologia* **35,** Suppl. 1, A4
38. Bullesbach, E. E., and Schwabe, C. (2005) *J. Biol. Chem.* **280,** 14586–14590
39. Bullesbach, E. E., and Schwabe, C. (1999) *Biochemistry* **38,** 3073–3078
40. Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N., and Vijayan, N. M. (1988) *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **319,** 369–456

Solution Structure and Novel Insights into the Determinants of the Receptor Specificity of Human Relaxin-3

K. Johan Rosengren, Feng Lin, Ross A. D. Bathgate, Geoffrey W. Tregear, Norelle L. Daly, John D. Wade and David J. Craik

J. Biol. Chem. 2006, 281:5845-5851.
doi: 10.1074/jbc.M511210200 originally published online December 19, 2005
