

# Interpretation of seemingly contradictory data: low NMR $S^2$ order parameters observed in helices and high NMR $S^2$ order parameters in disordered loops of the protein hGH at low pH

Lorna J. Smith,<sup>†</sup> Roya Athill,<sup>†</sup> Wilfred F. van Gunsteren,<sup>‡</sup> and Niels Hansen\*,<sup>¶</sup>

*Department of Chemistry, Inorganic Chemistry Laboratory, South Parks Road, Oxford,  
OX1 3QR, UK, Laboratory of Physical Chemistry, CH-8093 Zürich, Switzerland, and  
Institute of Thermodynamics and Thermal Process Engineering, D-70569 Stuttgart,  
Germany*

E-mail: hansen@itt.uni-stuttgart.de

Phone: +49 (0)711 685 66112. Fax: +49 (0)711 685 66140

## Abstract

At low pH human growth hormone (hGH) adopts a partially folded state in which the native helices are maintained but the long loop regions and side-chain packing becomes disordered. Some of the  $S^2$  order parameters for backbone N–H vectors derived from NMR relaxation measurements for the hGH at low pH initially seem contradictory. Three isolated residues (15, 20 and 171) in helices A and D exhibit low order-parameter values ( $< 0.5$ ) indicating flexibility while residue 143 in the centre of a long flexible

---

\*To whom correspondence should be addressed

<sup>†</sup>University of Oxford

<sup>‡</sup>Swiss Federal Institute of Technology, ETH

<sup>¶</sup>University of Stuttgart

loop region has a high order parameter (0.82). Using  $S^2$  order-parameter restraining MD simulations this paradox is solved. Low  $S^2$  values in helices are due to the presence of a mixture of  $3_{10}$ -helical and  $\alpha$ -helical hydrogen bonds. High  $S^2$  values in relatively disordered parts of a protein may be due to fluctuating networks of hydrogen bonds between backbone and side chains, which restrict the motion of N–H bond vectors.

## Introduction

Combining stability with flexibility within a single protein is an essential ingredient for its function such as catalysis, signalling or transport. Protein structure and stability may be elucidated by X-ray diffraction of crystals or by NMR spectroscopy in solution. The latter technique also allows to probe flexibility of parts of a protein through relaxation measurements. For example, squared order parameters  $S^2$  of the backbone N–H bond vectors can be derived from observed  $^{15}\text{N}$  NMR relaxation rates,<sup>1</sup> e.g. by applying the so-called "model-free" approach<sup>2,3</sup> i.c. model. Although such  $S^2$  order parameters are non-observable quantities, in contrast to the relaxation rates which are observable quantities from which  $S^2$  values can be derived using a model based on assumptions and approximations, they do reflect with a given uncertainty the variability of the direction of the N–H bond vectors with respect to the molecular frame. High  $S^2$  values indicate little variability and low  $S^2$  values indicate wide variability. Thus N–H  $S^2$  order parameters are used to probe flexibility of the protein backbone as function of residue number.

This view is confirmed by comparing the structural features of proteins for which a crystal structure could be derived from X-ray diffraction data and for which also NMR  $S^2$  order parameters could be derived from relaxation measurements of the same protein in aqueous solution. For example, the X-ray derived crystal structure of the human growth hormone (hGH)<sup>4</sup> shows four long helices A to D in a four-helix bundle fold with an up-up-down-down connectivity, see Figure 1. Figure S1 shows the N–H  $S^2$  order parameters for 108 of the 191 residues of this protein at pH = 7.0 (brown filled dots).<sup>5,6</sup> The four helices A (residues 9 –

34), B (residues 72 – 92, 94 – 100), C (residues 106 – 128) and D (residues 154 – 184) as observed in the crystal structure<sup>4</sup> do show high  $S^2$  values.

Under low pH conditions hGH undergoes a partial unfolding. No structures for the protein, derived from X-ray diffraction or NMR data, are available under low pH conditions. NMR chemical shift data and far-UV CD measurements suggest that the helical secondary structure of the protein is essentially unchanged from that observed at neutral pH.<sup>5,7,8</sup> However, near-UV CD data indicate that the side-chain packing becomes more disordered at low pH.<sup>7,8</sup> In addition, <sup>15</sup>N relaxation and hydrogen exchange studies show there is an increased flexibility for the protein as the pH is lowered in particular in the long loop regions.<sup>5</sup> The conformational variability of hGH at low pH is of interest, because the protein is stored in the pituitary gland in granules at an acidic pH,<sup>9–12</sup> and, on binding to its receptor, is exposed to broad-specificity proteases of lysosomes.<sup>13,14</sup> Similar behaviour had been identified for other four helix bundle proteins. For example, at low pH interleukin-4 is reported to populate a highly-ordered 'molten globule' state.<sup>15</sup>

Analysis of the N–H  $S^2$  order parameters for hGH at pH = 2.7, and comparison with the other experimental data available, identify some interesting features. Firstly, although the helical secondary structure is maintained under low pH conditions and most of the residues in the helices show high  $S^2$  values, some isolated N–H vectors in the middle of helices (residues 15 and 20 in helix A and residue 171 in helix D) display low  $S^2$  values ( $< 0.5$ ) indicating flexibility (Figure 2, black squared dots). In addition, although the N–H vectors of most residues in the loop between helices C and D have low order parameters ( $< 0.5$ ), residue 143 shows a high  $S^2$  value ( $> 0.8$ ), indicating structural stability. In the absence of flaws in the different experiments and procedures to determine the presence of helices and to derive  $S^2$  order parameters from relaxation rates, these observations pose a paradox.

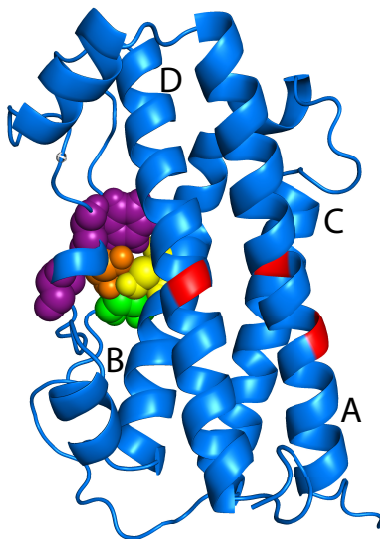


Figure 1: Ribbon representation of the structure of hGH as derived from X-ray diffraction data.<sup>4</sup> The four major helices A, B, C and D are indicated. Residues 15, 20 and 171 in red. Residue 143 in green. Side chains of Phe 54, Ser 57 (AB loop) in purple, of Ser 85 (B helix) in yellow, and of Ser 144 (CD loop) in orange.

## Solving the paradox using alternative experimental data?

To solve this paradox the literature was scanned for reports on proteins for which low backbone N–H  $S^2$  order parameters were reported within regions of helical secondary structure. For the B form of oxidised cytochrome  $b_5$  in 2M GdmCl a  $S^2$  value of  $0.41 \pm 0.12$  for Ser 64 in helix  $\alpha_5$  (residues 64 – 72) was reported.<sup>16</sup> For hen egg white lysozyme in trifluoroethanol residue 28 in helix II (residues 25 – 38) and residues 109, 111 and 112 in helix V (residues 104 – 114) show relatively low  $S^2$  values of  $0.65 \pm 0.07$ ,  $0.61 \pm 0.08$ ,  $0.71 \pm 0.11$ , and  $0.62 \pm 0.2$ , respectively.<sup>17</sup> The C3T mutant of the protein IL-4 at pH = 5 shows an  $S^2$  value of 0.61 for residue Phe 112 in helix D (residues 107 – 117).<sup>18</sup> In helix  $\alpha_1$  (residues 11 – 18) of thioredoxin residue 14 shows a lower than average order parameter.<sup>19</sup> Thus the observation of isolated residues with low  $S^2$  values in helical secondary structure found for hGH at low pH is not unique, but appears to be seen particularly in systems where the native state structure is slightly destabilised by, for example, mutation, low pH conditions or the addition of a low concentration of denaturant.

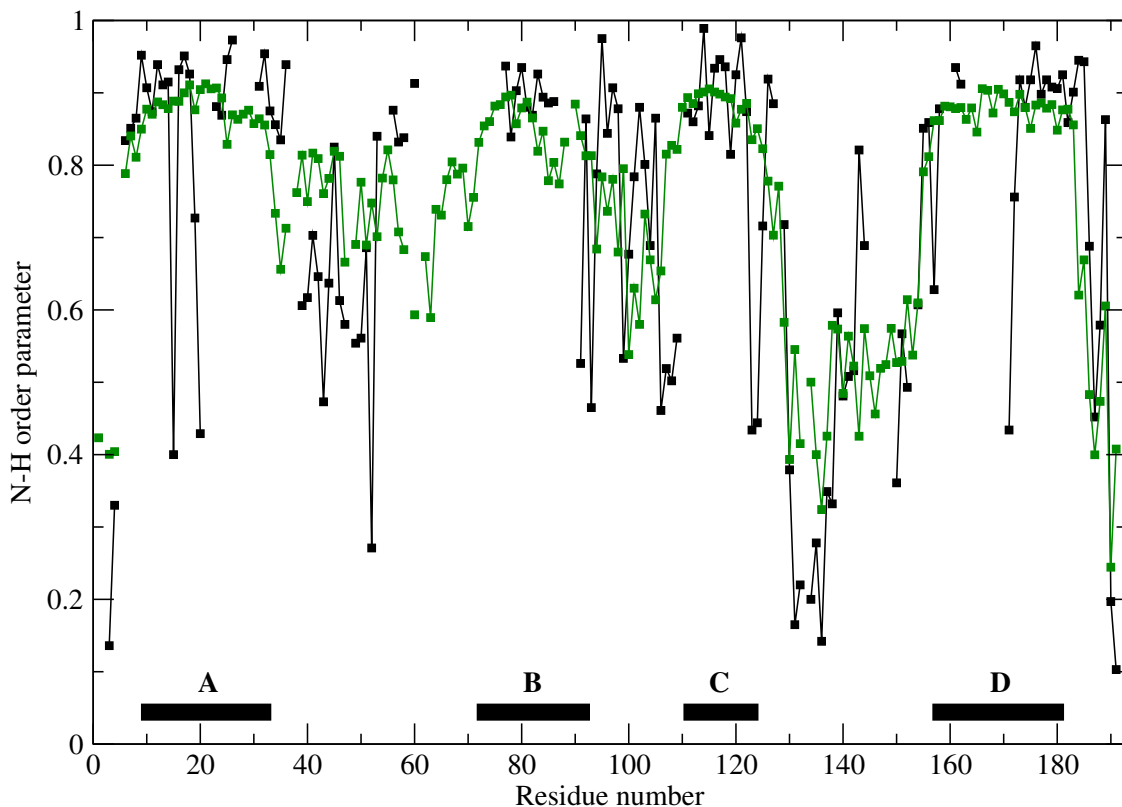


Figure 2: Backbone N – H bond vector  $S^2$  order parameters as function of residue number of hGH at pH = 2.7.  $S^2$  (relax) values derived from NMR relaxation data in black.<sup>5</sup>  $S^2$  (MD) values from unrestrained MD simulation in green. Only adjacent residues for which  $S^2$ -values are available are connected by the lines. The helices A to D are indicated by solid black bars.

The presence of residues with low  $S^2$  order parameter values in apparently persistent regions of helix may reflect different sensitivities of different experimental measurement techniques to the underlying structure. This can be investigated by calculating N–H  $S^2$  order parameters  $S^2(\delta)$  from measured protein backbone NMR chemical shift ( $\delta$ ) values using an empirical formula.<sup>20,21</sup> Figure 3 shows that the  $S^2(\delta)$  values calculated for hGH from the Random Coil Index (RCI) web-server<sup>20</sup> (brown squared dots, solid lines) generally display the same trend as the  $S^2$ (relax) values from the  $^{15}\text{N}$  relaxation data (black squared dots, solid lines), but with less variation. In particular the low  $S^2$ (relax) values for residues 15 and 20 in helix A and for residue 171 in helix D are not reproduced, and neither is the high  $S^2$ (relax) value for residue 143 in the loop between helices C and D. This is not unexpected,

because the empirical formula that connects  $S^2$  order-parameter values with chemical shift values is based on averaging of data from many proteins and thus has mean-field character.

In the BC loop the  $S^2(\text{relax})$  and  $S^2(\delta)$  values show differences (Figure 3). These probably reflect longer timescale micro-millisecond motions, which will average the chemical shifts resulting in lower predicted  $S^2(\delta)$  values. However, these motions will not cause lower  $S^2(\text{relax})$  values, but the data present in the BioMagRes data bank (BMRB entry 4689) identify relaxation-exchange contributions from the relaxation data analysis for residues 96-98 and 102 in this region (Figure S2).

We did search for other explanations for the presence of residues with low  $S^2$  order-parameter values in helices A and D. The errors in the order parameters given in the BioMagRes data bank (BMRB entry 4689) file are small for residues in the helices and for residues 15, 20 and 171 they are comparable to those for their neighbouring residues, which do not show low order-parameter values.

Hydrogen exchange data that are quoted in Ref. 5 hint at dynamics on long time scales. There are large variations in the hydrogen exchange protection factors observed, which are interpreted as reflecting the more disordered tertiary contacts in some parts of the structure. However, if the protection factors of adjacent residues in the helices are compared, which might experience similar levels of tertiary contact disorder, it is seen that residue 15 has a lower protection factor (63) than residues 14 and 16 (115 and 90, respectively), residue 20 has a lower protection factor (140) than residue 19 (250) and residue 171 has a lower protection factor (72) than residue 172 (437). For a mobile N-H group one would indeed expect lower hydrogen-exchange protection.

## Solving the paradox using simulation?

An alternative way to resolve the paradox is to resort to molecular dynamics (MD) simulations of hGH in aqueous solutions. At pH = 7 the profile of the backbone N-H  $S^2$  order

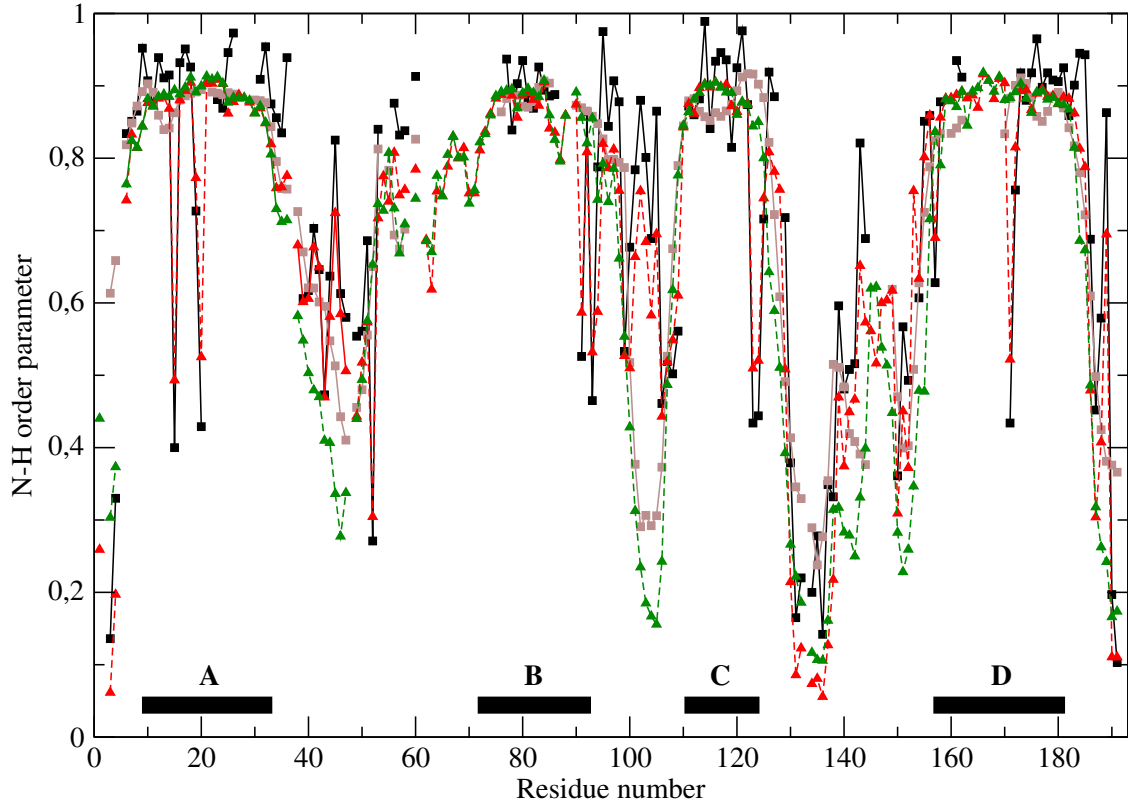


Figure 3: Backbone N – H bond vector  $S^2$  order parameters as function of residue number of hGH at pH = 2.7. Square dots and solid lines:  $S^2$  values derived from NMR data:  $S^2(\text{relax})$  values derived from NMR relaxation data<sup>5</sup> in black,  $S^2(\delta)$  values derived using an empirical relation<sup>20,21</sup> from NMR chemical shift data in brown. Triangles and dotted lines:  $S^2(\text{rMD})$  values from  $S^2$ -restraining MD simulations: Restraining towards  $S^2(\text{relax})$  values in red ( $S^2(\text{rMD}; \text{relax})$  simulation), restraining towards  $S^2(\delta)$  values in green ( $S^2(\text{rMD}; \delta)$  simulation).  $K^{\text{restr}} = 600 \text{ kJ mol}^{-1}$  and  $\tau = 20 \text{ ps}$ . Only adjacent residues for which  $S^2$ -values are available are connected by the lines. The helices A to D are indicated by solid black bars.

parameter (brown filled dots in Figure S1) is rather well reproduced by MD simulation using the GROMOS 54A7 force field<sup>22</sup> and simulation software<sup>23,24</sup> (red filled dots in Figure S1). At pH 7, there are aggregation issues which may affect the  $S^2$  values. Therefore, the MD simulation at pH 7 is not considered further. At pH 2.7 the NMR spectra<sup>5</sup> are concentration independent. In addition, dynamic light scattering measurements performed under the same conditions are consistent with hGH being monomeric. At low pH hGH is highly charged (+24e) which will make aggregation less likely.

Lowering the pH to 2.7 acidic side chains become protonated, but this does not strongly affect the  $S^2(\text{MD})$  profile (green squares in Figure 2). Only at the end of helix B and in the loop connecting helices C and D the  $S^2(\text{MD})$  values are significantly lower at pH = 2.7 than at pH = 7. The low  $S^2(\text{relax})$  values for residues 15 and 20 in helix A and residue 171 in helix D are not reproduced by  $S^2(\text{MD})$ . The relatively high  $S^2(\text{relax})$  value of residue 143 in the loop between helices C and D is also not reproduced by  $S^2(\text{MD})$ . To interpret results of MD simulations appropriately, the quality of the calculated properties must be evaluated. This depends on (1) the degrees of freedom simulated, (2) the accuracy of the molecular model, interaction function or force field, (3) the equations of motion, integration scheme or other method used to sample degrees of freedom as well as the degree of sampling, (4) the boundary conditions, (5) the simulation software, and (6) how the software is used. If an experimentally measured value of a particular property, such as an  $S^2$  order parameter, is not reproduced in a simulation, this can have many reasons, e.g. a force-field deficiency or insufficient sampling of particular degrees of freedom. These deficiencies can be redressed by restraining the motion of the molecule such that an experimental target value is on average reproduced in the simulation. Such time-averaged restraining enhances the configurational sampling by forcing the molecule to surmount barriers that would without restraining only be surmounted rarely, that is, on long time scales. In other words, the time-averaged restraining induces long time motion to occur on a shorter time scale thereby widening the configurational sampling. By restraining the motion of N-H vectors to reproduce low  $S^2$  order-parameter values, the local motion that without restraining would cover a long time scale is sped up to the time scale of the memory relaxation time used in the restraining. This means that the long-time configurational distribution of the N-H vectors is sampled on a short time scale. This is one of the advantages of restraining while accounting for averaging.<sup>25</sup> In view of this it was decided to apply  $S^2$  order-parameter restraining<sup>26</sup> to the backbone N-H bond vectors in the MD simulation using either the  $S^2(\text{relax})$  values or the  $S^2(\delta)$  values as target values  $S_0^2$  for the restraining. This, of course, led to a much better reproduction of the  $S^2(\text{relax})$  and



$S^2(\delta)$  profiles by the respective  $S^2(\text{rMD})$  profiles (red and green triangles in Figure 3). Since the low  $S^2$  values in helices A and D were reproduced in the MD simulation restraining the  $S^2$  values towards the  $S^2(\text{relax})$  values while retaining the helical structure, analysis of this simulation would allow an explanation of the paradox.

## Reconciling low $S^2$ values with helical structure

Table 1 shows the  $S^2$  order parameters and hydrogen-bond population of the N–H groups at pH 2.7 of the three residues, 15, 20 and 171 in helices A and D that show low  $S^2(\text{relax})$  values together with those quantites for the neighbouring residues. The low  $S^2$  value is correlated with the occurrence of a  $3_{10}$ -helical hydrogen bond of the N–H moiety. Switching between  $\alpha$ -helical and  $3_{10}$ -helical hydrogen bonding, the N–H bond vector lying in the peptide plane changes direction lowering the  $S^2$  value. The overall helical structure is maintained during the switching. The helix is locally changing its pitch.

One may ask whether such a switching between  $\alpha$ - and  $3_{10}$ -helical hydrogen bonding would be detected based on NMR measurements. In Table 2 the HN(i) – HN(i+1) and HA(i-3) – HN(i) distances, for which NOE intensity can be identified for helical fragments, are shown as averages over the entire trajectory and over trajectory structures displaying either  $\alpha$ -helical or  $3_{10}$ -helical hydrogen bonding involving the HN(i) moiety. The differences in the HA(i-3) – HN(i) distance are too small to reliably distinguish between  $\alpha$ - and  $3_{10}$ -helix. The HN(i) – HN(i+1) distance is for the  $3_{10}$ -helical structure lower than for the  $\alpha$ -helical structure, but the difference is not large enough to allow for a distinction between these two types of helix based on standard NOE data obtained by conversion of NOE cross peak intensity into a distance range based on calibration of a few intra-residue distances and NOE intensities. If a mixture of  $\alpha$ -helical and  $3_{10}$ -helical structure is present, as for the residues of Table 2, no conclusion about their populations can be drawn from the NOE data. These observations imply that the model and force field used to derive protein structure from NMR

spectroscopic data will determine the type of helix, not the measured NMR data.

In the MD simulation without  $S^2$  order-parameter restraining  $3_{10}$ -helical hydrogen bonds were not observed for residues 15, 19, 20 and 171. When starting such an unrestrained simulation from a structure with  $3_{10}$ -helical hydrogen bonds at the mentioned residues, these hydrogen bonds switch to  $\alpha$ -helical ones. This indicates that the GROMOS 54A7 force field parameter set favours  $\alpha$ -helical over  $3_{10}$ -helical hydrogen bonding. This is in accordance with the  $\alpha$ -helical and  $3_{10}$ -helical propensities as observed for protein structures deposited in the Protein Data Bank. However, the free-energy difference between  $\alpha$ -helical and  $3_{10}$ -helical structures may be slightly overestimated by the 54A7 force-field parameter set, which minor force-field deficiency is easily (low  $K^{\text{restr}}$  value) corrected by the  $S^2$  order-parameter restraining. Test simulations with low  $S^2$  order parameter target values for the N-H vector of residues other than 15, 19, 20 and 171 but within helices all show a mixture of  $\alpha$ -helical and  $3_{10}$ -helical hydrogen bonding in order to attain low  $S^2$  values, confirming the observed correlation between low  $S^2$  value and  $\alpha$ - and  $3_{10}$ -helical mixture within helical fragments.

## Reconciling high $S^2$ with disordered loop structure

Table 3 shows the percentage hydrogen bonding between the hydrogen-bond donors and acceptors in the spatial surroundings of the NH group of Tyr 143. The direction of the N-H vector is maintained by a fluctuating set of hydrogen bonds between atoms of neighbouring side chains and some backbone carbonyl oxygens. Thus the structure of this part of the CD loop is not a single one, but changes while keeping the variation of the direction of the N-H bond vector of residue 143 limited, leading to a relatively high  $S^2$  value for this residue. In a follow-up study of ref. 5, the temperature dependence of  $S^2(\text{relax})$  values for hGH was reported.<sup>27</sup> The  $S^2(\text{relax})$  value for residue 143 is significantly reduced as the temperature is increased (see Figure 2 in ref. 27). This would be compatible with the MD simulation results. The fluctuating hydrogen-bonds around residue 143 are likely to be easily lost at

higher temperatures, as they involve side-chain atoms. The behaviour of the CD loop region in the MD simulation presents a model for how the overall tertiary structure of a protein can be maintained in a partially folded state even when the side-chain packing is disordered.

## Variation of model parameters

The results discussed in the previous section depend on the way the  $S^2$  order-parameter values were derived from measured values of observables and on the way the  $S^2$ -restraining MD simulation was carried out. The  $S^2(\delta)$  values derived empirically from chemical shift data do not show the low  $S^2$  values for residues 15, 20 and 171, and thus the MD simulation using the  $S^2(\delta)$  values as target for  $S^2$  order-parameter restraining do not show significant  $3_{10}$ -helical structure around these residues, 0 % for residue 15, 1 % for residue 20 and 2 % for residue 171 (Table 1). In a NMR spectroscopic study of the stability of hGH as a function of temperature,<sup>27</sup>  $S^2$  order parameters for residues 15, 19, 20 and 171 were not reported, which may be due to the inability of the authors to explain the compatibility of low  $S^2$  values with helical structure.

To evaluate the influence of the  $S^2$ -restraining parameters  $K^{\text{restr}}$  (weight) and  $\tau$  (averaging time) on the simulated results, these parameters were varied,  $K^{\text{restr}} = 300, 400$  or  $600 \text{ kJ mol}^{-1}$  and  $\tau = 20$  or  $200 \text{ ps}$ , for both sets of target data  $S^2(\text{relax})$  and  $S^2(\delta)$ . These variations from the values  $K^{\text{restr}} = 600 \text{ kJ mol}^{-1}$  and  $\tau = 20 \text{ ps}$  do not lead to significantly different results (see Supporting Information).

## Implications for biomolecular function

One of the interesting features of hGH is that conformational flexibility of the loop regions seem to enable the protein to adapt its local structure such as to bind to a variety of different receptor proteins.<sup>28</sup> This appears to be particularly important for residues 38 to 47 of hGH, which are involved in receptor binding, and form a helix in some structures of hGH

in complex with growth hormone or prolactin receptors.<sup>28,29</sup> The internal motions of hGH relevant for receptor binding are likely to occur over a range of time scales including those that are much longer than the fast ps-ns motions considered in this work. However, fast ps-ns motions have been recognised to influence both the thermodynamics and kinetics of protein functions including binding processes.<sup>1,30-32</sup> It is therefore interesting that the analysis of the hGH MD simulations showed that some residues in the region 38-46 also adopted  $\alpha$ - as well as  $3_{10}$ -helical structure, as was observed for helices A and D. Here, however, the helical population is significantly reduced under low pH conditions (27-56 % and 10-22 % population of  $\alpha$ -helical and  $3_{10}$ -helical hydrogen bonds, respectively). This is in accord with the experimental finding that the Phe 44 – Leu 45 peptide bond is susceptible to proteolysis at low pH.<sup>33</sup> It is proposed that the partial unfolding of this helical region may occur when the hormone is exposed to an acidic environment in the cell and that the proteolysis would generate two physiologically relevant fragments.

# Conclusion

Using different measurement techniques to obtain structural information on biomolecules, it sometimes happens that seemingly contradictory information is obtained. Such a paradox can be due to approximations and simplifications employed when deriving molecular structure, a non-observable quantity, from measured values of observable quantities.<sup>25</sup> Such paradoxes can be resolved by applying MD simulation, possibly with restraining to a set of values of particular quantities that are to be satisfied to allow for a reliable interpretation of the simulated results.<sup>34,35</sup> Here we showed that low  $S^2$  order-parameter values for backbone N–H bond vectors are perfectly compatible with helical local structure of a protein. They indicate the presence of both  $\alpha$ - and  $3_{10}$ -helical hydrogen bonds involving the N–H group. The lower the  $S^2$  value the more  $3_{10}$ -helical hydrogen bonding is replacing the dominant  $\alpha$ -helical hydrogen bonding. Inversely, high  $S^2$  order parameter values for backbone N–H bond vectors

can be found in less ordered, non-helical, parts of the protein. These can be due to a set of fluctuating hydrogen bonds between backbone and side chains of neighbouring residues. These results are independent of details of the  $S^2$  value restraining in the simulations.

Interpretation of molecular properties based on values of non-observable quantities such as molecular structure and order parameters, which are derived from measured values of observable quantities, such as X-ray diffraction intensities and NMR relaxation rates, respectively, may be hampered by the assumptions and approximations made when converting the measured data into values of non-observable quantities.<sup>25,36</sup> This was illustrated by the comparison of  $S^2$  order parameters for N-H bond vectors obtained from observed NMR relaxation rates for hGH with  $S^2$  values derived from chemical shift data using an empirical relation. The differences observed between both sets of  $S^2$  values led to a different structural interpretation of the  $S^2$  values, which constitutes a warning for a too simplistic interpretation of experimental data. Finally we note that neither NOE intensities nor chemical shifts measured by NMR are able to distinguish between  $\alpha$ -helical and  $3_{10}$ -helical structure of protein fragments. This implies that their presence in protein structures derived from such NMR data is determined by the molecular model and force field used in the structure determination or refinement, not by the measured data.

## Materials and methods

The MD simulations were carried out using the GROMOS biomolecular simulation software<sup>23,24</sup> and the GROMOS 54A7 force-field parameter set.<sup>22</sup> The starting coordinates for hGH in the simulations were taken from the X-ray structures of hGH in complex with the extracellular domain of its receptor with Protein Data Bank code 3HHR.<sup>4</sup> The coordinates of residues 149-153 in an exposed loop, of the C-terminal residue 191 and some of the side-chain atoms of residues 135-148 are missing from this structure. These were added using Superlooper<sup>37</sup> and Swiss PDB Viewer.<sup>38</sup> For the pH 7.0 simulation the Asp and Glu side

chains in the protein were unprotonated and the histidine side chains were singly protonated at N $\delta$ 1 for His 18 and at N $\epsilon$ 2 for His 21 and 151. For the pH 2.7 simulations the Asp and Glu side chains and the protein C-terminus were protonated and the histidine side chains were doubly protonated. For both the pH 7.0 and pH 2.7 simulations, the protein was solvated in a rectangular box and minimum image periodic boundary conditions were applied. The minimum solute-box wall distance was set to 1.1 nm giving 20423 and 21021 simple point charge (SPC) water<sup>39</sup> molecules at pH 7.0 and pH 2.7 respectively. To achieve overall neutrality of the system five sodium ions were added to the pH 7.0 simulation and 24 chloride ions were added to the pH 2.7 simulation.

For each system an equilibration scheme comprising five short simulations at temperatures of 60 K, 120 K, 180 K, 240 K and 305 K was used at constant volume. These periods of equilibration were each 20 ps long in the simulation at pH 7.0 and 100 ps long in the simulation at pH 2.7. During the first four of the equilibration periods the solute atoms were harmonically restrained to their positions in the initial structures with force constants of 25000, 2500, 250 and 25 kJ mol<sup>-1</sup> nm<sup>-2</sup> at temperatures of 60 K, 120 K, 180 K and 240 K, respectively. Following this equilibration, simulations were performed at a temperature of 305 K and a pressure of 1 atm using the weak-coupling algorithm,<sup>40</sup> with relaxation times of  $\tau_T = 0.1$  ps and  $\tau_p = 0.5$  ps and an isothermal compressibility of  $4.575 \times 10^{-4}$  (kJ mol<sup>-1</sup> nm<sup>-3</sup>)<sup>-1</sup>. Solute and solvent were separately coupled to the heat bath. The SHAKE algorithm<sup>41</sup> was used to constrain bond lengths of the solutes and the rigid geometry of the solvent molecules, with a relative geometric tolerance of  $10^{-4}$  allowing for an integration time step of 2 fs. The centre of mass motion of the system was removed every 1000 time steps. Non-bonded interactions were calculated using a triple-range cut-off scheme<sup>42</sup> applied on the basis of distances between charge group centers,<sup>43</sup> with cutoff radii of 0.8 nm and 1.4 nm. Interactions within 0.8 nm were evaluated every time step and intermediate interactions were updated every fifth time step. To account for the influence of the dielectric medium outside the 1.4 nm cutoff sphere, a reaction-field force<sup>44,45</sup> with a

dielectric permittivity  $\varepsilon$  of 61 was used.<sup>46</sup>

In addition to the unrestrained simulations, simulations at pH 2.7 were also run with  $^1\text{H}$ - $^{15}\text{N}$  order parameter restraints applied.<sup>26</sup> Two different sets of  $^1\text{H}$ - $^{15}\text{N}$  order parameter restraints were used. Set  $S^2(\text{relax})$  used the experimental order parameters determined at pH 2.7 (BMRB entry 4689; ref. 5). Set  $S^2(\delta)$  used order parameters calculated from the chemical shift data (backbone  $\text{C}\alpha$ ,  $\text{CO}$ ,  $\text{N}$ ,  $\text{H}\alpha$ ,  $\text{NH}$  chemical shifts) for hGH at pH 2.7 (BMRB entry 4689; ref. 5) using the Random Coil Index (RCI) webserver.<sup>20</sup> In both cases, order-parameter values which were greater than 0.9 were set to 0.9. For each order-parameter set 20 ns simulations were run with a force constants of 300, 400 and 600  $\text{kJ mol}^{-1}$ , memory relaxation times of 20 and 200 ps and a flat bottom<sup>26</sup> of  $\Delta S^2 = 0.05$ .

Analysis was performed with the GROMOS++ suite of analysis programs,<sup>47</sup> using coordinate and energy trajectories written to disk every 5 ps over a simulation period of 20 ns. Order parameters were calculated using an averaging time window of 1 ns. Hydrogen bonds were identified according to a geometric criterion: a hydrogen bond was assumed to exist if the hydrogen-acceptor distance is smaller than 0.25 nm and the donor-hydrogen-acceptor angle is larger than  $135^\circ$ .

Table 1: Backbone N–H bond vector  $S^2$  order parameters and populations (%) of NH–CO hydrogen bonds with low  $S^2$  values in helices A and D of hGH at pH 2.7<sup>a</sup>.

backbone							
NH(i) – CO(j)	helix type	$S^2$ (relax)	$S^2$ (rMD; relax)	% H-bond in rMD(relax)	$S^2(\delta)$	$S^2(\text{rMD}; \delta)$	% H-bond in rMD( $\delta$ )
H – bond i – j							
14 – 10	$\alpha$	0.91	0.87	96	0.84	0.89	98
15 – 11	$\alpha$	0.40	0.49	66	0.86	0.89	99
15 – 12	$3_{10}$	0.40	0.49	20	0.86	0.89	0
16 – 12	$\alpha$	0.93	0.88	94	0.88	0.89	93
17 – 13	$\alpha$	0.95	0.89	97	0.89	0.90	97
18 – 14	$\alpha$	0.93	0.90	97	0.89	0.91	99
19 – 15	$\alpha$	0.73	0.77	75	0.89	0.89	92
19 – 16	$3_{10}$	0.73	0.77	7	0.89	0.89	2
20 – 16	$\alpha$	0.43	0.53	52	0.89	0.90	90
20 – 17	$3_{10}$	0.43	0.53	20	0.89	0.90	1
21 – 17	$\alpha$	–	0.90	99	–	0.91	99
170 – 166	$\alpha$	–	0.90	99	0.83	0.88	97
171 – 167	$\alpha$	0.43	0.52	54	0.88	0.88	87
171 – 168	$3_{10}$	0.43	0.52	24	0.88	0.88	2
172 – 168	$\alpha$	0.76	0.82	88	0.90	0.89	95
173 – 169	$\alpha$	0.92	0.89	96	0.91	0.90	98

<sup>a</sup> i, j: residue numbers. Simulated values from restrained MD simulations at pH = 2.7 with  $K^{\text{restr}} = 600 \text{ kJ mol}^{-1}$  and  $\tau = 20 \text{ ps}$ . A hydrogen bond is present when the distance H – O is smaller than 0.25 nm and the angle N – H – O is larger than  $135^\circ$ .  $S^2(\text{relax})$  obtained from NMR relaxation data.  $S^2(\delta)$  obtained from NMR chemical shift data. Restraining in MD either to the former or the latter data.



Table 2: Average HN(i) – HN(i+1) and HA(i-3) – HN(i) distances (nm) and populations (between parentheses in %) for particular N-H atoms in helices A and D of hGH over the entire trajectory or when forming  $\alpha$ -helical or  $3_{10}$ -helical hydrogen bonds, respectively in the  $S^2$  restrained MD simulations at pH = 2.7 using either  $S^2(\text{relax})$  or  $S^2(\delta)$  as target values<sup>a</sup>.

	atom-atom distance in rMD(relax)			atom-atom distance in rMD( $\delta$ )		
	traj	$\alpha$ -helix	$3_{10}$ -helix	traj	$\alpha$ -helix	$3_{10}$ -helix
15HN – 16HN	0.290	0.305 (65.8)	0.228 (19.8)	0.299	0.299 (98.7)	0.249 (0.3)
12HA – 15HN	0.355	0.353	0.342	0.345	0.345	0.342
19HN – 20HN	0.297	0.303 (75.2)	0.259 (7.5)	0.295	0.298 (91.8)	0.255 (2.4)
16HA – 19HN	0.357	0.355	0.358	0.345	0.345	0.347
20HN – 21HN	0.287	0.303 (52.0)	0.226 (19.8)	0.294	0.295 (90.5)	0.237 (0.7)
17HA – 20HN	0.350	0.347	0.340	0.336	0.336	0.335
171HN – 172HN	0.285	0.303 (54.4)	0.231 (24.1)	0.291	0.294 (87.1)	0.241 (1.6)
168HA – 171HN	0.370	0.368	0.360	0.361	0.359	0.363

<sup>a</sup> Simulated values from restrained MD simulations with  $K^{\text{restr}} = 600 \text{ kJ mol}^{-1}$  and  $\tau = 20 \text{ ps}$ .

Table 3: Populations (%) of hydrogen bonds around residue 143 of hGH from  $S^2$  restrained MD simulations at pH = 2.7 using either  $S^2(\text{relax})$  or  $S^2(\delta)$  as target values<sup>a</sup>.

donor – acceptor pair		% H-bond in rMD (relax)	% H-bond in rMD ( $\delta$ )
Ser 57 OGH	- Gln 141 O	21	3
	- Thr 142 O	-	25
	- Tyr 143 OH	16	-
Gln 84 NE2H	- Tyr 143 OH	-	17
Ser 85 OGH	- Tyr 143 O	6	-
Thr 142 OG1H	- Ser 85 OG	19	1
Tyr 143 OHH	- Ser 57 OG	10	-
Tyr 143 OHH	- Gln 84 OE1	-	28
Tyr 143 NH	- Gln 141 O	1	14
Tyr 143 NH	- Ser 85 O	7	-
Ser 144 OGH	- Ser 85 O	24	-
Ser 144 NH	- Phe 54 O	-	41

<sup>a</sup> Simulated values from restrained MD simulations at pH = 2.7 with  $K^{\text{restr}} = 600 \text{ kJ mol}^{-1}$  and  $\tau = 20 \text{ ps}$ . A hydrogen bond is present when the distance H – O is smaller than 0.25 nm and the angle N – H – O is larger than 135°. Only hydrogen bonds which have a population of greater than 5% in one of the simulations are included.

## Acknowledgement

LJS and RA would like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out some of this work. <http://dx.doi.org/10.5281/zenodo.22558>. WFvG thanks the Swiss National Science Foundation, grant number 200020-137827, and the European Research Council, grant number 228076, for financial support. NH thanks the German Research Foundation (DFG) for financial support within the Cluster of Excellence in Simulation Technology (EXC 310/2) at the University of Stuttgart.

## Supporting Information Available

Figure showing  $S^2(\text{relax})$  values derived from NMR relaxation data at pH = 7 and  $S^2(\text{MD})$  values from unrestrained MD simulations. Figure showing the positions of residues in the hGH structure for which an  $R_{\text{ex}}$  term was included in the analysis of the  $^{15}\text{N}$  relaxation data. Tabulated  $S^2$  order parameter values at pH 2.7 derived from NMR relaxation and chemical shift data, respectively, as well as obtained from unrestrained and restrained MD simulation using different restraining parameters.

## References

- (1) Jarymowycz, V. A.; Stone, M. J. *Chem. Rev.* **2006**, *106*, 1624–1671.
- (2) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4559.
- (3) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559–4570.
- (4) De Vos, A. M.; Ultsch, M.; Kossiakoff, A. A. *Science* **1992**, *255*, 306–312.
- (5) Kasimova, M. R.; Kristensen, S. M.; Howe, P. W.; Christensen, T.; Matthiesen, F.; Petersen, J.; Sørensen, H. H.; Led, J. J. *J. Mol. Biol.* **2002**, *318*, 679–695.
- (6) Jensen, M. R.; Kristensen, S. M.; Keeler, C.; Christensen, H. E. M.; Hodsdon, M. E.; Led, J. J. *Proteins, Struct. Funct. Bioinf.* **2008**, *73*, 161–172.
- (7) Kasimova, M. R.; Milstein, S. J.; Freire, E. *J. Mol. Biol.* **1998**, *277*, 409 – 418.
- (8) DeFelippis, M. R.; Kilcomons, M. A.; Lents, M. P.; Youngman, K. M.; Havel, H. A. *Biochim. Biophys. Acta* **1995**, *1247*, 35–45.
- (9) Dannies, P. S. *Endocr. Rev.* **1999**, *20*, 3–21.
- (10) Dannies, P. S. *Mol. Cell. Endocrinol.* **2001**, *177*, 87–93.
- (11) Lee, M. S.; Zhu, Y. L.; Chang, J. E.; Dannies, P. S. *J. Biol. Chem.* **2001**, *276*, 715–721.

- (12) Weisz, O. A. *Int. Rev. Cytol.* **2003**, *226*, 259–319.
- (13) Authier, F.; Posner, B. I.; Bergeron, J. J. *FEBS Lett.* **1996**, *389*, 55–60.
- (14) Van Kerkhof, P.; Strous, G. *Biochem. Soc. Trans.* **2001**, *29*, 488–493.
- (15) Redfield, C.; Smith, R. A. G.; Dobson, C. M. *Nature Struct. Biol.* **1994**, *1*, 23–29.
- (16) Arnesano, F.; Banci, L.; Bertini, I.; Koulougliotis, D.; Monti, A. *Biochemistry* **2000**, *39*, 7117–7130.
- (17) Buck, M.; Schwalbe, H.; Dobson, C. M. *J. Mol. Biol.* **1996**, *257*, 669–683.
- (18) Vaz, D. C. The role of disulphides on the structure, packing and stability of human Interleukin-4. Ph.D. thesis, University of Coimbra, Portugal, 2010.
- (19) Stone, M. J.; Chandrasekhar, K.; Holmgren, A.; Wright, P. E.; Dyson, H. J. *Biochemistry* **1993**, *32*, 426–435.
- (20) Berjanskii, M. V.; Wishart, D. S. *J. Am. Chem. Soc.* **2005**, *127*, 14970–14971.
- (21) Berjanskii, M. V.; Wishart, D. S. *J. Biomol. NMR* **2008**, *40*, 31–48.
- (22) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. *Eur. Biophys. J.* **2011**, *40*, 843–856.
- (23) van Gunsteren, W. F. 2015; <http://www.gromos.net>.
- (24) Schmid, N.; Allison, J. R.; Dolenc, J.; Eichenberger, A. P.; Kunz, A.-P. E.; van Gunsteren, W. F. *J. Biomol. NMR* **2011**, *51*, 265–281.
- (25) van Gunsteren, W. F.; Allison, J. R.; Daura, X.; Dolenc, J.; Hansen, N.; Mark, A. E.; Oostenbrink, C.; Rusu, V. H.; Smith, L. J. *Angew. Chem. Int. Ed.* **2016**, *55*, 15990–16010, *Angew. Chem.* 2016, *128*, 16222–16244.

- (26) Hansen, N.; Heller, F.; Schmid, N.; van Gunsteren, W. F. *J. Biomol. NMR* **2014**, *60*, 169–187.
- (27) Vinther, J. M.; Kristensen, S. M.; Led, J. J. *J. Am. Chem. Soc.* **2010**, *133*, 271–278.
- (28) Kossiakoff, A. A.; Somers, W.; Ultsch, M.; Andow, K.; Muller, Y. A.; De Vos, A. M. *Protein Sci.* **1994**, *3*, 1697–1705.
- (29) Pearce, K. H.; Ultsch, M. H.; Kelley, R. F.; De Vos, A. M.; Wells, J. A. *Biochemistry* **1996**, *35*, 10300–10307.
- (30) Wand, A. J. *Nature Struct. Biol.* **2001**, *8*, 926–921.
- (31) Karplus, M.; Kuriyan, J. *Proc. Nat. Acad. Sci. USA* **2005**, *102*, 6679–6685.
- (32) Kovermann, M.; Rogne, P.; Wolf-Watz, M. *Q. Rev. Biophys.* **2016**, *49*, e6.
- (33) Spolaore, B.; Polverino de Laureto, P.; Zambonin, M.; Fontana, A. *Biochemistry* **2004**, *43*, 6576–6586.
- (34) Trzesniak, D.; Glättli, A.; Jaun, B.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **2005**, *127*, 14320–14329.
- (35) Trzesniak, D.; Van Gunsteren, W. F. *Protein Sci.* **2006**, *15*, 2544–2551.
- (36) van Gunsteren, W. F.; Dolenc, J.; Mark, A. E. *Curr. Opin. Struct. Biol.* **2008**, *18*, 149–153.
- (37) Hildebrand, P. W.; Goede, A.; Bauer, R. A.; Gruening, B.; Ismer, J.; Michalsky, E.; Preissner, R. *Nucleic Acids Res.* **2009**, *37*, W571–W547.
- (38) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714–2723.
- (39) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Inter-molecular Forces*; Pullmann, B., Ed.; Reidel: Dordrecht, the Netherlands, 1981; pp 331–342.

- (40) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (41) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (42) Berendsen, H. J. C.; van Gunsteren, W. F.; Zwinderman, H. R. J.; Geurtsen, R. G. *Ann. N. Y. Acad. Sci.* **1986**, *482*, 269–285.
- (43) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag AG an der ETH Zürich: Zürich, Groningen, 1996.
- (44) Barker, J. A.; Watts, R. O. *Mol. Phys.* **1973**, *26*, 789–792.
- (45) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- (46) Heinz, T. N.; van Gunsteren, W. F.; Hünenberger, P. H. *J. Chem. Phys.* **2001**, *115*, 1125–1136.
- (47) Eichenberger, A. P.; Allison, J. R.; Dolenc, J.; Geerke, D. P.; Horta, B. A. C.; Meier, K.; Oostenbrink, C.; Schmid, N.; Steiner, D.; Wang, D.; van Gunsteren, W. F. *J. Chem. Theory Comput.* **2011**, *7*, 3379–3390.

## Graphical TOC Entry

