# Transient Hydrogen Bonds Identified on the Surface of the NMR Solution Structure of Hirudin†

Thomas Szyperski,‡ Walfrido Antuch,‡ Martin Schick,‡ Andreas Betz,§ Stuart R. Stone,§ and Kurt Wüthrich*,‡

*Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg, CH-8093 Zürich, Switzerland, and Department of Haematology, University of Cambridge, MRC Centre, Hills Road, Cambridge CB2 2QH, U.K.*

*Received April 15, 1994*®

ABSTRACT: Recombinant desulfatohirudin retains largely the thrombin-inhibitory activity of natural hirudin from *Hirudo medicinalis* and causes at most minimal immune response in humans. With regard to potential pharmaceutical applications it is of interest to further investigate the structural basis of hirudin functions. In this paper transient hydrogen bonds between backbone amide protons and side-chain carboxylates on the protein surface of desulfatohirudin (variant 1) have been identified using two-dimensional $^1$H NMR experiments and site-directed mutagenesis. The analysis of pH titration curves measured with NMR enabled the determination of the p$K$ values of all 13 carboxylates, and downfield shifts larger than 0.2 ppm arising from weak bonding interactions with carboxylates were observed for the amide protons of Gly 25, Ser 32, Glu 35, and Cys 39. For these backbone amide protons virtually identical titration parameters were observed in intact desulfatohirudin and the mutant, truncated hirudin(1–51), demonstrating that the hydrogen bond acceptors are located in the N-terminal polypeptide segment 1–51. The hydrogen bonds Gly 25 NH–Glu 43 $\delta$COO$^-$, Ser 32 NH–Glu 35 $\delta$COO$^-$, Glu 35 NH–Asp 33 $\gamma$COO$^-$, Glu 35 NH–Glu 35 $\delta$COO$^-$, and Cys 39 NH–Glu 17 $\delta$COO$^-$ were identified by considering spatial proximity in the NMR solution structure of hirudin(1–51), and comparing the p$K$ values for the amide protons and the carboxylates in desulfatohirudin and the mutants hirudin(E43Q), hirudin(E35Q), hirudin(D33N) and hirudin(E17A). Comparative structure calculations with and without distance constraints for these hydrogen bonds showed that although they are all compatible with the NMR solution structure, these hydrogen bonds are transient dynamic features of the protein surface which, with the sole exception of Cys 39 NH–Glu 17 $\delta$COO$^-$, would not have been detected in a conventional NMR structure determination. Of special interest is the clear-cut information obtained on the fact that the lifetimes of the dynamic "bifurcated" hydrogen-bonding interactions of the amide proton of Glu 35 are in the millisecond time range or shorter.

Hirudin variant 1 (HV1;[1] Scharf et al., 1989) is a small protein of 65 amino acid residues occurring in the salivary glands of the leech *Hirudo medicinalis*. It is the most potent known inhibitor of the blood clotting enzyme thrombin, which is a serine protease that plays a central role in the pathology of thrombotic diseases (Johnson et al., 1989). HV1 contains a sulfated tyrosyl residue in position 63 (Badgy et al., 1976; Dodt et al., 1984), while recombinant desulfatohirudin lacks this posttranslational modification. Nevertheless, desulfatohirudin binds to thrombin nearly as tightly as natural hirudin (Stone & Hofsteenge, 1986) and causes no or only minimal immune response in humans (Close et al., 1994), which makes it an attractive polypetide for therapeutic applications in cardiovascular medicine (Lent, 1986; Märki & Wallis, 1990). Notwithstanding current limitations of our knowledge on structural features that determine immune response (Langone, 1989), future deeper insights will undoubtedly depend on the availability of detailed descriptions of the protein surfaces which constitute the epitopes that are recognized by the immune system. Further investigation of the protein surface of hirudin is thus clearly of direct interest, and in this paper we describe the characterization of transient hydrogen bonds on the surface of hirudin in aqueous solution by NMR techniques.

The NMR solution structure of desulfatohirudin contains a globular amino-terminal domain of residues 1–48 and a flexibly disordered carboxy-terminal tail of residues 49–65 (Folkers et al., 1989; Haruyama & Wüthrich, 1989). Desulfatohirudin contains 13 carboxylates (eight glutamates, four aspartates, and the C-terminal carboxylate), of which seven are located in the C-terminal segment 53–65. In earlier work (Haruyama et al., 1989), pH titration shifts of amide protons were detected which must be due to hydrogen-bonding interactions with some of these carboxylate groups (Bundi & Wüthrich, 1977, 1979). As was previously demonstrated in model peptides (Bundi & Wüthrich, 1977, 1979) as well as in proteins (Ebina & Wüthrich, 1984; O'Connell et al., 1993; Steinmetz et al., 1988), amide proton titration shifts may manifest transient hydrogen bonds with carboxylate groups that would not be identified in a convential NMR structure determination based on measurements of NOEs and spin–spin coupling constants. In desulfatohirudin, there is *a priori*

† This work was supported by the Schweizerischer Nationalfonds (Project 31.32033.91), MRC (S.R.S.), and by a Bundesstipendium to W.A.

‡ Eidgenössische Technische Hochschule-Hönggerberg.

§ University of Cambridge.

® Abstract published in *Advance ACS Abstracts*, July 1, 1994.

[1] Abbreviations: HV1, hirudin variant 1; desulfatohirudin, recombinant hirudin HV1; hirudin(1–51), N-terminal 51-residue polypeptide segment of desulfatohirudin; hirudin(E35Q), desulfatohirudin with Glu 35 replaced by Gln; hirudin(E17A), desulfatohirudin with Glu 17 replaced by Ala; hirudin(D33N), desulfatohirudin with Asp 33 replaced by Asn; hirudin(E43Q), desulfatohirudin with Glu 43 replaced by Gln; NMR, nuclear magnetic resonance; 2D, two-dimensional; P. COSY, 2D purged correlation spectroscopy; TOCSY, 2D total correlation spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, 2D nuclear Overhauser enhancement spectroscopy; TPPI, time-proportional phase incrementation; TSP, [2,2,3,3-$^2H_4$]trimethylsilyl)propionate; ppm, parts per million; $\delta_{HA}$, chemical shift in ppm of the fully protonated state; $\delta_{A^-}$, chemical shift in ppm of the fully deprotonated state; $\delta$(pH), experimental chemical shift in ppm at a specified pH value; p$K$, negative logarithm of the acid dissociation constant.

0006-2960/94/0433-9303$04.50/0 

Table 1: Survey of 2D $^1$H NMR Spectra Recorded ($T$ = 22 °C)

| experiment[a] | $\tau_m$[b] | data matrix[c] | $\tau_{1max}$[d] | $\tau_{2max}$[d] | $\tau_{tot}$[e] | resolution[f] |
|---|---|---|---|---|---|---|
| | | 6 mM desulfatohirudin/1.77, 2.14, 2.60, 3.30, 3.73, 4.10, 4.41, 4.77, 5.12, 5.75, 6.75[g] | | | | |
| NOESY | 80 | 230 × 2048 | 38 | 336 | 12 | 6.0/1.5 |
| | | 6 mM desulfatohirudin/2.60, 2.80, 3.50, 3.95, 4.35, 5.30[g] | | | | |
| clean-TOCSY | 100 | 320 × 1024 | 57 | 182 | 8 | 5.5/2.7 |
| | | 3 mM hirudin(E17A)/2.61, 3.21, 3.73, 4.13, 5.62, 6.48[g] | | | | |
| NOESY | 100 | 256 × 1024 | 42 | 168 | 9 | 6.0/3.0 |
| | | 3 mM hirudin(E35Q)/2.02, 2.97, 3.55, 3.73, 4.23, 4.74, 5.06, 6.44[g] | | | | |
| NOESY | 100 | 195 × 1024 | 33 | 170 | 9 | 6.0/3.0 |
| | | 3 mM hirudin(E43Q)/2.17, 3.86, 4.26, 4.62, 4.97, 6.43[g] | | | | |
| NOESY | 100 | 210 × 1024 | 37 | 180 | 9 | 5.6/2.7 |
| | | 3 mM hirudin(D33N)/2.15, 2.97, 3.44, 3.99, 4.45, 5.02, 6.34[g] | | | | |
| NOESY | 100 | 500 × 1024 | 88 | 180 | 11.5 | 5.6/2.7 |
| | | 2 mM hirudin(1–51)/2.60, 2.80, 3.50, 3.95, 4.35, 5.30[g] | | | | |
| P.COSY | | 256 × 2048 | 42 | 336 | 13.5 | 6.0/1.5 |

[a] References for the experimental schemes used are as follows: P. COSY, Marion and Bax (1988); clean-TOCSY, Griesinger et al. (1988); NOESY, Anil-Kumar et al. (1980). [b] Mixing time in milliseconds. [c] Size of the acquired data matrices in complex points. [d] Maximal $t_1$ and $t_2$ values in milliseconds. [e] Approximate total measuring time per spectrum in hours. [f] Spectral resolution along $\omega_1/\omega_2$ after zero-filling, in Hz/point. [g] Sample/pH values.

the intriguing question to be answered whether such hydrogen-bonding contacts are exclusively between atom groups within the globular domain of residues 1–48 or also between amide protons in the structured domain and the numerous carboxylates in the C-terminal flexible tail. From the earlier studies with desulfatohirudin (Haruyama et al., 1989), some long-range interactions with residues near the C-terminus were implicated, but these conclusions had to be based on limited experimental data. In particular, the three-dimensional structure available for intact desulfatohirudin (Folkers et al., 1989; Haruyama & Wüthrich, 1989) was not of very good quality because of interference of NMR signals from the flexible tail with the data collection for the globular domain. Spectral overlap also prevented measurements of some of the 13 carboxylate pH titration curves over a sufficiently large range for a reliable determination of the p$K$ values. In the meantime, a high-quality NMR structure was determined for the truncated hirudin(1–51) (Szyperski et al., 1992). In addition, as described in this paper, a sufficient selection of mutant proteins with single-residue replacements was prepared to resolve all ambiguities in the identification of amide proton–carboxylate hydrogen bonds, so that a complete description of this type of transient surface side-chain–backbone interactions can now be presented.

## MATERIALS AND METHODS

Desulfatohirudin variant 1 (Scharf et al., 1989) was used as it was given to us by Ciba-Geigy AG, Basel, Switzerland (Meyhack et al., 1987; Grossenbacher et al., 1987). Hirudin(1–51) and the mutants hirudin(E43Q), hirudin(E17A), hirudin(D33N), and hirudin(E35Q) were produced with recombinant DNA techniques as described by Dennis et al. (1990) and Braun et al. (1988).

*NMR Spectroscopy.* 2D $^1$H NMR experiments were recorded on a Bruker AM500 spectrometer at 500 MHz and at 22 °C in the pure-phase absorption mode using TPPI (Marion & Wüthrich, 1983). The protein solutions were prepared in a mixed solvent of 90% $H_2O$ and 10% $^2H_2O$, and the pH value was adjusted by adding small amounts of HCl and NaOH. The proton chemical shifts were calibrated relative to TSP at 0.00 ppm. The spectra were processed on a Bruker X32 workstation using the program UXNMR. The residual water signal after preirradiation was further reduced using the convolution method of Marion et al. (1989). Before Fourier transformation, the time domain data were zero-filled and multiplied with shifted sine-bell windows (DeMarco & Wüthrich, 1976). The spectra were baseline-corrected using third-order polynomials. Further experimental details are given in Table 1.

*Identification of Hydrogen Bonds with pH Titration Experiments.* The protocol to identify hydrogen-bonding interactions between amide protons and carboxylates (Bundi & Wüthrich, 1977, 1979) is based on changing the pH from acidic to neutral. The sign of the pH variation in chemical shift indicates when a certain amide proton is involved in such a hydrogen bond; *i.e.*, upfield shifts are mediated via covalent bonds for the amide protons of Asp residues and the C-terminal amino acid, and downfield shifts are observed for amide protons that are hydrogen bonded to carboxylates. Obviously, the pH titration curves of the amide protons and the intrinsic titration curves of the interacting carboxylates, as observed on the $\gamma$-protons of Glu, the $\beta$-protons of Asp, or the $\alpha$-proton of the C-terminal residue, must correspond to identical p$K$ values. Provided that electrostatic interferences with other ionizable groups are negligible, the pH titration curves for interacting amide protons and carboxylate groups can be parametrized with a single p$K$ value. In a globular protein, identity of the p$K$ values for an amide proton and the corresponding carboxylate can be used as a necessary condition for identifying a hydrogen-bonding interaction, provided that the global conformation does not vary with pH over the range used. Having a sufficiently well resolved structural description of the protein at hand, it may then be possible to uniquely identify pairs of carboxylate groups and amide protons in spatial proximity that have identical p$K$ values. Residual ambiguities due to, for example, degeneracy of two or multiple carboxylate p$K$ values in the same protein can be resolved by site-directed mutagenesis. Any point mutation that replaces a carboxylate without changing the global protein structure leads to the disappearance of all amide proton downfield shifts arising from hydrogen bonding with this carboxylate and thus provides the desired identification of the acceptor residue.

*Quantitative Data Analysis.* Titration parameters were obtained by nonlinear least-squares fits of the one-proton titration curve (eq 1),

$$\delta\,(\mathrm{pH}) = \frac{\delta_{HA} + \delta_{A^-} 10^{\mathrm{pH}-\mathrm{p}K}}{1 + 10^{\mathrm{pH}-\mathrm{p}K}} \qquad (1)$$

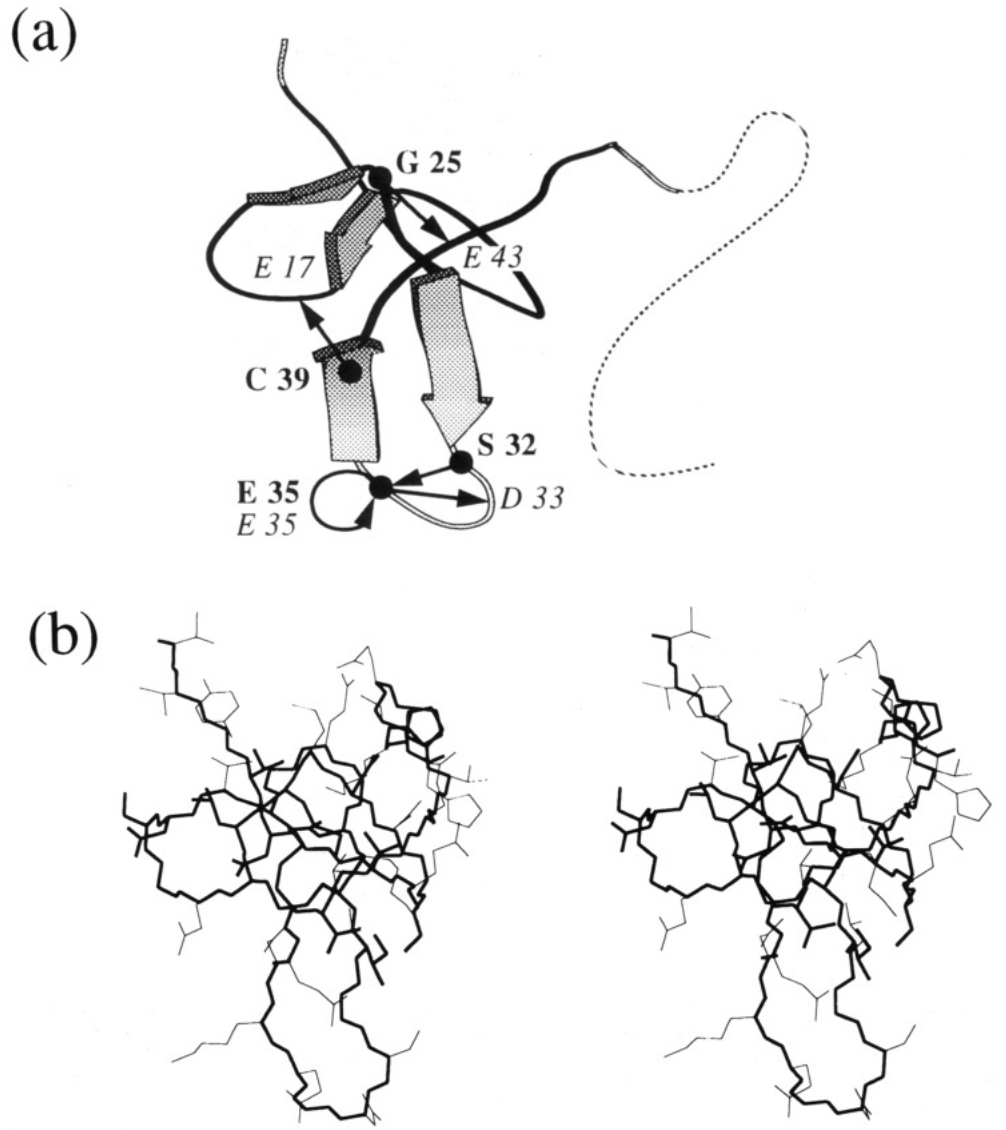


FIGURE 1: (a) Ribbon drawing of the NMR solution structure of desulfatohirudin variant 1 generated with the program MOLSCRIPT (Kraulis, 1991) using the conformer of hirudin(1–51) with the lowest DIANA target function value (Szyperski et al., 1992) and the knowledge that segment 52–65 is flexibly disordered in solution (Folkers et al., 1989; Haruyama & Wüthrich, 1989). The arrowed ribbons indicate position and direction of the $\beta$-sheet strands. A black rope represents the well-structured loop areas. Residues 1, 2, 32–36, and 49–51, which are only poorly constrained by the NMR data for hirudin(1–51), are indicated by a white rope, and residues 52–65, by a dotted line. The locations of backbone amide protons with $\Delta\delta(H^N) > 0.2$ ppm are identified by black circles and bold lettering, using the one-letter amino acid code and the sequence position. Arrows point to the $C^\alpha$ positions of the hydrogen bond acceptors identified by italic lettering. (b) Stereoview of an all-heavy-atom presentation of the hirudin(1–51) NMR conformer with the lowest DIANA target function value (Szyperski et al., 1992) in the same orientation as in (a). The backbone and the 19 best-defined side chains (amino acid residues 4–7, 9, 12, 14–16, 19, 20, 22, 26, 28, 29, 39, 40, 46, 48) are displayed with bold lines, and the other side chains, with thin lines.

to the experimental chemical shift, $\delta$(pH), where $\delta_{HA}$ is the chemical shift in the acidic pH limit, while $\delta_{A^-}$ represents the chemical shift in the basic pH limit. Equation 1 can be rearranged to

$$\log\left[\frac{\delta\,(\mathrm{pH}) - \delta_{HA}}{\delta_{A^-} - \delta\,(\mathrm{pH})}\right] = \mathrm{pH} - \mathrm{p}K \qquad (2)$$

In order to obtain the standard deviation, $\sigma$, of the p$K$ value, we introduced $\delta_{HA}$ and $\delta_{A^-}$ as obtained from the nonlinear fit of eq 1 into eq 2 and performed a linear fit of eq 2 to the experimental data. The overlap of the 99% confidence intervals of the p$K$ values measured for the amide proton and the interacting carboxylate was used as a necessary condition for the formation of a hydrogen bond.

Calculations of the complete three-dimensional structure using the input data set described by Szyperski et al. (1992) supplemented by distance constraints for the presently identified hydrogen bonds were performed for hirudin(1–51) with the distance geometry program DIANA (Güntert et al., 1991). The distance between an amide proton and the oxygen atom of the interacting carboxylate was restricted to 1.8–2.0 Å, and the distance between the nitrogen atom of the amide group and the oxygen atom of the carboxylate was restricted to 2.7–3.0 Å (Williamson et al., 1985).

## RESULTS

The molecular structure of desulfatohirudin (Figure 1), which represents the scaffold for the present investigation of transient features of the protein surface, includes a globular domain, with two antiparallel $\beta$-sheets (residues 14–16 and 20–22, and 27–31 and 36–40), three reverse turns (residues 8–11, type II; 17–20, type II′; and 23–26, type II), and a short stretch of a polyproline helix II (Sasisekharan, 1959) (residues 46–48), and the flexibly disordered segment 49–65. The NMR structure determination of hirudin(1–51) (Szyperski et al.,

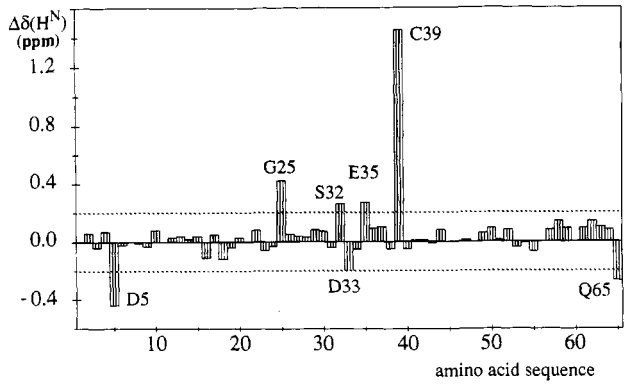


FIGURE 2: Plot of the titration shifts in ppm, $\Delta\delta(H^N)$, measured for the backbone amide protons of desulfatohirudin *versus* the amino acid sequence, which is VVYTD CTESG QNLCL CEGSN VCGQG NKCIL GSDGE KNQCV TGEGT PKPQS HNDGD FEEIP EEYLQ. $\Delta\delta(H^N) = \delta(H^N, pH = 1.77) - \delta(H^N, pH = 6.75)$, and amino acid residues with $\Delta\delta(H^N) > |0.2|$ ppm are labeled.

1992) revealed that removal of the carboxy-terminal polypeptide segment 52–65 does not noticeably affect the static and dynamic properties of the amino-terminal domain; the loop of residues 31–36 and the N-terminal dipeptide segment are flexibly disordered in both the intact and the truncated protein. For the present project, it is of particular relevancy that 12 of the 13 acidic groups in recombinant hirudin exhibit flexibly disordered conformations that allow for transient interactions with backbone amide protons, the only well-structured carboxylate being that of Asp 5.

*NMR Experiments.* The identification of the hydrogen-bonding partners described below depended importantly on detailed investigations of the pH dependence of their NMR parameters over the pH range 2–6 (Bundi & Wüthrich, 1977, 1979). An overview of all NMR experiments performed is afforded by Table 1. Phase-sensitive NOESY spectra (Anil-Kumar et al., 1980) at variable pH values were recorded for desulfatohirudin and for its mutants hirudin(E43Q), hirudin-(E17A), hirudin(D33N), and hirudin(E35Q). Due to spectral overlap it was not possible to monitor the chemical shifts of the γ-protons of Glu 62 over a wide range of pH values in these spectra; therefore an additional series of clean-TOCSY spectra (Griesinger et al., 1988) were acquired for desulfatohirudin. The pH titration shifts of the amide proton resonances in hirudin(1–51) were followed by P. COSY spectra (Marion et al., 1988). In total, 50 2D NMR experiments were performed, with a total measurement time of about 21 days for the present investigation (Table 1). Systematic shifts of proton resonances that might occur due to variations in ionic strength or protein concentration were assessed from the observation of the proton resonances of the methyl groups of Val 21, Leu 15, and Leu 30. These chemical shifts remained constant within 0.01 ppm over the entire pH range from 2 to 6, so that global conformational rearrangements over this pH range could be excluded. This was further confirmed by the analysis of numerous medium- and long-range NOEs.

*Identification of Amide Protons Acting as Hydrogen Bond Donors.* Figure 2 presents a survey of the chemical shift variations of the backbone amide protons of recombinant hirudin, $\Delta\delta(H^N)$, when going from an acidic (pH = 1.77) to a neutral milieu (pH = 6.75). Large downfield shifts ($\Delta\delta(H^N) > 0.2$ ppm) indicative of through-space interactions with carboxylates are observed for the amide protons of Gly 25 ($\Delta\delta(H^N) = 0.42$ ppm), Ser 32 ($\Delta\delta(H^N) = 0.26$ ppm), Glu 35 ($\Delta\delta(H^N) = 0.27$ ppm), and Cys 39 ($\Delta\delta(H^N) = 1.45$ ppm). The corresponding titration parameters that were obtained

Table 2: Titration Parameters for Amide Protons with Titration Shifts Exceeding 0.2 ppm and for Side-Chain Protons of Glu and Asp in Desulfatohirudin and Hirudin Mutants at $T = 22$ °C

| residue[a] | proton | protein[b] | $\delta_{HA}$[c] (ppm) | $\Delta\delta$[c] (ppm) | p$K \pm 2.5\sigma$[d] |
|---|---|---|---|---|---|
| Asp 5 | $\beta^1$ | | 2.80 | −0.32 | 4.25 ± 0.09 |
| | $\beta^2$ | | 2.91 | −0.26 | 4.28 ± 0.13 |
| | $H^N$ | | 9.02 | −0.44 | 4.32 ± 0.09 |
| Glu 8 | γ | | 2.48 | −0.18 | 4.28 ± 0.17 |
| **Glu 17** | $\gamma^2$ | | **2.48** | **−0.23** | **3.79 ± 0.17** |
| **Gly 25** | $H^N$ | | **8.72** | **0.42** | **4.25 ± 0.07** |
| | $H^N$ | **hirudin(1–51)** | **8.69** | **0.47** | **4.29 ± 0.20** |
| | $H^N$ | **hirudin(E43Q)** | **8.86** | **−0.01** | |
| **Ser 32** | $H^N$ | | **8.18** | **0.26** | **4.31 ± 0.07** |
| | $H^N$ | **hirudin(1–51)** | **8.20** | **0.24** | **4.33 ± 0.20** |
| | $H^N$ | **hirudin(E35Q)** | **8.18** | **−0.02** | |
| **Asp 33** | $\beta^1$ | | **2.85** | **−0.20** | **4.24 ± 0.17** |
| | $\beta^2$ | | **3.14** | **−0.29** | **4.12 ± 0.05** |
| | $H^N$ | | 9.10 | −0.20 | 3.93 ± 0.10 |
| | $\beta^2$ | **hirudin(E35Q)** | **3.16** | **−0.30** | **3.86 ± 0.10** |
| **Glu 35** | $H^N$ | | **7.64** | **0.27** | **4.00 ± 0.10** |
| | $H^N$ | **hirudin(1–51)** | **7.65** | **0.27** | **3.87 ± 0.17** |
| | $H^N$ | **hirudin(D33N)** | **7.63** | **0.14** | **4.32 ± 0.12** |
| **Gln 35** | $H^N$ | **hirudin(E35Q)** | **7.66** | **0.15** | **3.91 ± 0.11** |
| **Glu 35** | $\gamma^1$ | | **2.60** | **−0.31** | **4.32 ± 0.07** |
| | $\gamma^2$ | | **2.65** | **−0.21** | **4.36 ± 0.07** |
| | $\gamma^1$ | **hirudin(D33N)** | **2.58** | **−0.24** | **4.26 ± 0.10** |
| | $\gamma^2$ | **hirudin(D33N)** | **2.63** | **−0.15** | **4.36 ± 0.12** |
| **Cys 39** | $H^N$ | | **8.80** | **1.45** | **3.76 ± 0.07** |
| | $H^N$ | **hirudin(1–51)** | **8.79** | **1.41** | **3.78 ± 0.15** |
| | $H^N$ | **hirudin(E17A)** | **9.10** | **−0.04** | |
| **Glu 43** | $\gamma^1$ | | **2.53** | **−0.29** | **4.25 ± 0.11** |
| | $\gamma^2$ | | **2.62** | **−0.23** | **4.24 ± 0.07** |
| Asp 53 | $\beta^1$ | | 2.90 | −0.25 | 3.78 ± 0.15 |
| | $\beta^2$ | | 2.97 | −0.26 | 3.79 ± 0.15 |
| Asp 55 | $\beta^1$ | | 2.78 | −0.25 | 4.07 ± 0.07 |
| | $\beta^2$ | | 2.86 | −0.25 | 4.17 ± 0.06 |
| Glu 57 | γ | | 2.36 | −0.17 | 4.63 ± 0.10 |
| Glu 58 | $\beta^2$ | | 2.07 | −0.09 | 4.69 ± 0.05 |
| Glu 61 | γ | | 2.48 | −0.19 | 4.48 ± 0.15 |
| Glu 62 | $\gamma^1$ | | 2.30 | −0.13 | 4.53 ± 0.12 |
| | $\gamma^2$ | | 2.34 | −0.14 | 4.51 ± 0.16 |
| Gln 65 | α | | 4.29 | −0.16 | 4.01 ± 0.10 |
| | $H^N$ | | 8.16 | −0.36 | 3.78 ± 0.15 |

[a] Bold letting identifies residues involved in NH–$COO^-$ hydrogen bonds. [b] No entry means desulfatohirudin. [c] $\Delta\delta = \delta_{A^-} - \delta_{HA}$, with $\delta_{A^-}$ and $\delta_{HA}$ as obtained from a nonlinear least-squares fit of eq 1 to the experimental data. [d] $\sigma$ is the standard deviation for the determination of the p$K$ value (see text).

by nonlinear least-squares fits of eq 1 to the experimental data (Table 2) coincide closely with those reported by Haruyama et al. (1989). The spatial locations of the backbone amide protons in the polypeptide chain are indicated in Figure 1a. The downfield shifts of all side-chain amide protons were found to be smaller than 0.1 ppm, suggesting that none of the side-chain amide moieties is involved in a significantly populated hydrogen bond with a carboxylate. Two additional intriguing features in the data of Figure 2 are that there are numerous residues in the flexible segment 49–65 with $\Delta\delta(H^N) = 0.1$–$0.15$ ppm, indicating formation of significantly populated transient hydrogen bonds, and that only three residues (Asp 5, Asp 33, and Gln 65) show the expected large intrinsic high-field shifts (Bundi & Wüthrich, 1977), whereas for Asp 53 and Asp 55 these are presumably offset by downfield shifts arising from hydrogen-bonding interactions.

*Identification of Carboxylates Acting as Hydrogen Bond Acceptors.* The titration parameters of all 13 carboxylate groups in desulfatohirudin are given in Table 2. In addition, Figure 3 affords a visual display of the distribution of p$K$ values and their 99% confidence intervals, showing that the p$K$ values are well dispersed. The small standard deviations reflect that all titration curves exhibit sigmoidal shapes (see Figures 4–7). Thus, electrostatic interference between different carboxylates, which would typically lead to nonsigmoidal

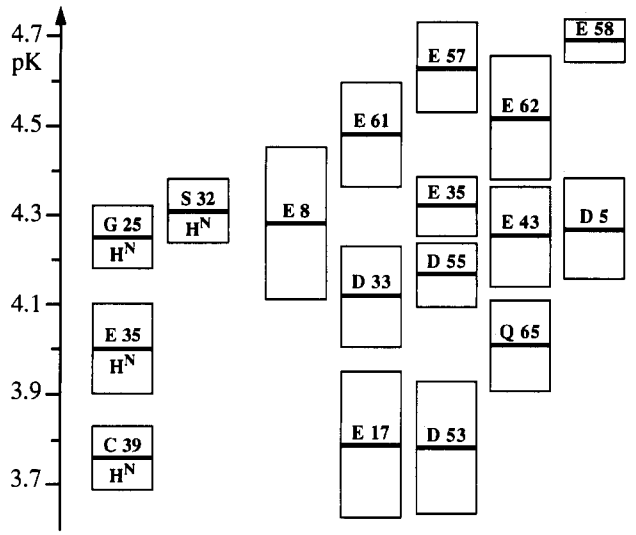


FIGURE 3: Graphical representation of the p$K$ values of backbone amide protons with $\Delta\delta(H^N) > 0.2$ ppm (see Figure 2 and Table 2) and the p$K$ values of the 13 carboxylates in desulfatohirudin (thick horizontal lines). For carboxylates where the p$K$ value was determined by the observation of two protons, the mean value is displayed (see Table 2). The 99% confidence intervals (corresponding to $2.5\sigma$, where $\sigma$ is the standard deviation for the determination of the p$K$ values) are represented by boxes.

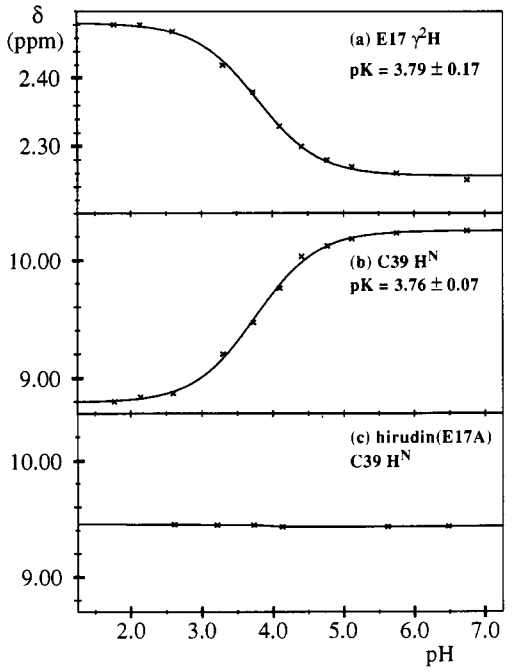


FIGURE 4: Plots of chemical shifts *versus* pH of (a) the $\gamma^2$ proton of Glu 17 in desulfatohirudin, (b) the amide proton of Cys 39 in desulfatohirudin, and (c) the amide proton of Cys 39 in hirudin(E17A). The curves were determined by a nonlinear least-squares fit of eq 1 to the experimental data. The p$K$ values are also given.

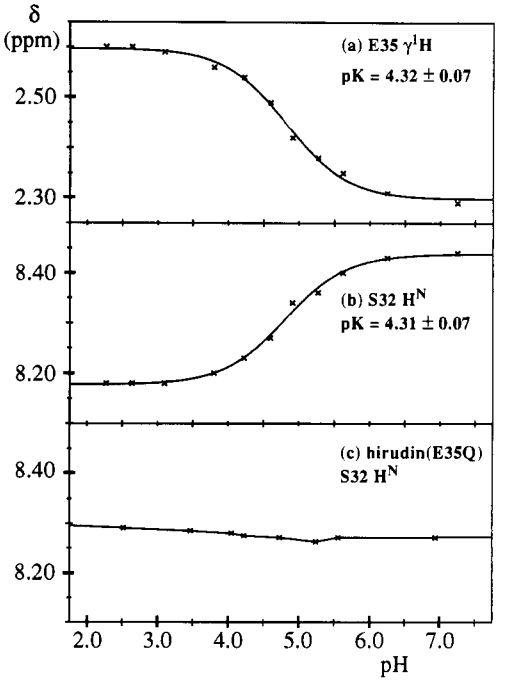


FIGURE 5: Same as Figure 4 for (a) the $\gamma^1$ proton of Glu 35 in desulfatohirudin, (b) the amide proton of Ser 32 in desulfatohirudin, and (c) the amide proton of Ser 32 in hirudin(E35Q).

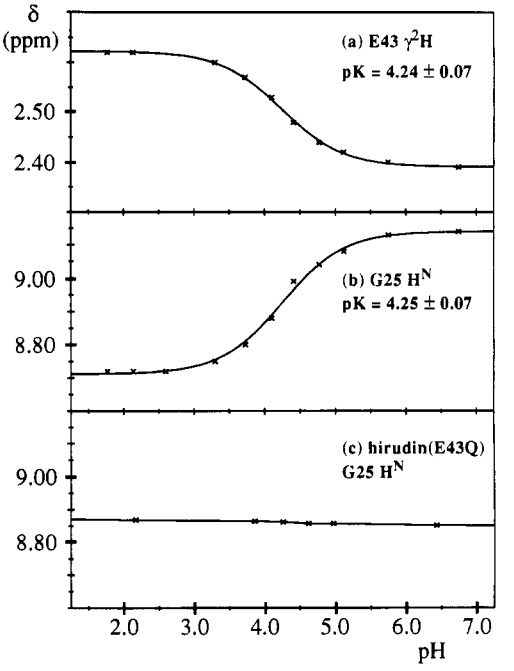


FIGURE 6: Same as Figure 4 for (a) the $\gamma^2$ proton of Glu 43 in desulfatohirudin, (b) the amide proton of Gly 25 in desulfatohirudin, and (c) the amide proton of Gly 25 in hirudin(E43Q).

titration curves, appears not to play a significant role, suggesting that the negative charges of the carboxylates are well screened due to the high dielectric constant of the aqueous solvent milieu with the elevated ionic strength arising from the high protein concentration used. Now that a complete data set is available, Figure 3 illustrates why unambiguous identification of the hydrogen bond acceptors cannot be based solely on the criterion that the 99% confidence intervals of the p$K$ values of the interacting groups must overlap (Haruyama et al., 1989). Since the supplementary criterion of compatibility with the NMR structure of desulfatohirudin could not fully clarify the assignments, mainly because of the flexible nature of most carboxylate-bearing residues (see above and Figure 1b), we resorted to the use of mutant proteins tailored to resolve the remaining uncertainties.

A first important result was that the titration parameters of the backbone amide protons of Gly 25, Ser 32, Glu 35, and Cys 39 were virtually unchanged in hirudin(1–51) when compared with the values obtained for desulfatohirudin (Table 2), so that all carboxylates located in the flexible C-terminal tail comprising residues 49–65 can be excluded as potential hydrogen bond acceptors. Conclusive individual hydrogen bond assignments were then obtained by comparison of wild-type desulfatohirudin with mutants in which the most likely

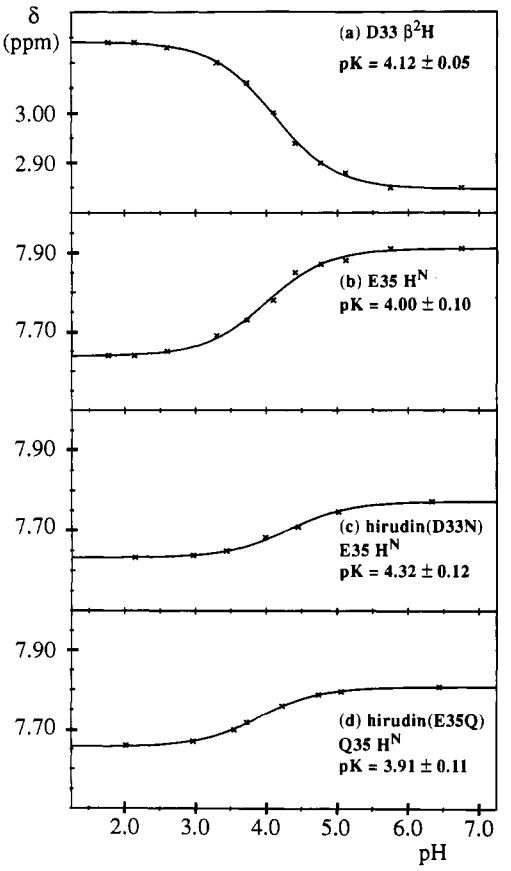


FIGURE 7: Same as Figure 4 for (a) the $\beta^2$ proton of Asp 33 in desulfatohirudin, (b) the amide proton of Glu 35 in desulfatohirudin, (c) the amide proton of Glu 35 in hirudin(D33N), and (d) the amide proton of Gln 35 in hirudin(E35Q).

acceptor groups were substituted by the corresponding amides or, in the case of Glu 17, by Ala. In all these mutants, chemical shifts of nonlabile protons and long-range NOEs were monitored over the pH range used to ascertain that observed variations in the titration parameters could not be attributed to global conformational rearrangements. Thus the interactions Cys 39 NH–Glu 17 $\delta COO^-$ (Figure 4), Ser 32 NH–Glu 35 $\delta COO^-$ (Figure 5), and Gly 25 NH–Glu 43 $\delta COO^-$ (Figure 6) were clearly identified from the complete disappearence of the pH dependence of the chemical shifts in the mutant proteins (Table 2). A more complex situation is encountered for the backbone amide proton of Glu 35, since a residual downfield titration shift of $\Delta\delta(H^N) = 0.1$ ppm was observed in hirudin(D33N) (Figure 7c), with a p$K$ value close to that of the carboxylate of Glu 35 (Table 2). The downfield shift of the amide proton of Gln 35 in hirudin(E35Q) was also reduced to $\Delta\delta(H^N) = 0.14$ ppm (Figure 7d). Thus, the amide proton of Glu 35 interacts with both the carboxylate of Asp 33 and the carboxylate of its own side chain, while in turn the carboxylate of Glu 35 acts as a hydrogen bond acceptor for both the amide proton of Ser 32 and its own amide proton.

*Structure Calculations with Hydrogen Bond Constraints.* In order to test whether the presently identified hydrogen bonds are compatible with the input data set of 580 NOE upper distance constraints and 109 dihedral angle constraints used previously for the structure determination of hirudin(1–51), we performed new distance geometry calculations using the program DIANA (Güntert et al., 1991) with supplementary distance constraints (Williamson et al., 1985) for the interresidual hydrogen bonds Gly 25 NH–Glu 43 $\delta COO^-$, Ser 32 NH–Glu 35 $\delta COO^-$, Glu 35 NH–Asp 33 $\gamma COO^-$, and Cys 39 NH–Glu 17 $\delta COO^-$. Otherwise, the same protocol was used as reported by Szyperski et al. (1992). Compared to results obtained using the input without hydrogen bond constraints, the group of 20 conformers with the lowest residual DIANA target function values exhibited only slightly increased residual constraint violations, with target function values between 0.6 and 0.8 Å$^2$, as compared with 0.08–0.17 Å$^2$ (Szyperski et al., 1992). This clearly demonstrates that the hydrogen-bonding constraints are consistent with the constraints derived from $^1$H–$^1$H NOEs and scalar spin–spin coupling constants. A visual impression of the hydrogen bonds is afforded by the stereo pictures in Figure 8, which were taken from the DIANA conformer with the lowest target function value (the structures of Figure 8 may be compared with the corresponding segments in the all-heavy-atom representation of hirudin(1–51) in Figure 1b, which was computed without supplementary hydrogen bond constraints). Clearly, the static view of Figure 8 is an oversimplified description of the protein surface, since we now know, for example, that the side chain of Glu 35 interacts with both the amide proton of Ser 32 and its own amide proton. Since only a single averaged resonance is seen for each of these two amide protons as well as for the Glu 35 side chain, these hydrogen bonds must be formed and cleaved in a rapid, dynamic equilibrium, with lifetimes of the order of milliseconds or shorter (Wüthrich, 1986).

## DISCUSSION

*Locations of the Amide Proton–Carboxylate Hydrogen Bonds.* A survey of the locations of the hydrogen bonds in the desulfatohirudin structure is presented in Figure 1a. The hydrogen bond involving the carboxylate of Glu 17 and the backbone amide proton of Cys 39 connects the type II′ turn comprising residues 17–20 with the second strand of the $\beta$-sheet containing residues 27–31 and 36–40. These two secondary structure elements are locally and globally well defined in the NMR solution structure of hirudin(1–51) (see Figures 2 and 3 in Szyperski et al. (1992)), imply that the backbone atoms of Glu 17 and Cys 39 are in fixed relative orientations. Hence, the entropic expense for this hydrogen bond formation is restricted to the loss of the degrees of freedom about the side-chain torsion angles of Glu 17. The following three arguments indicate that this is a highly populated hydrogen bond. First, the downfield shift measured for the amide protons of Cys 39 (Figures 2 and 4) is more than 2 times larger than the largest previously observed values in peptides (Bundi & Wüthrich, 1977, 1979) or proteins (*e.g.*, Ebina and Wüthrich (1984), O'Connell et al. (1992), and Steinmetz et al. (1988)). Second, this hydrogen bond has been identified in 12 out of the 20 energy-refined DIANA conformers used to represent the NMR solution structure of hirudin(1–51) that was calculated without supplementary hydrogen bond constraints using the common distance and bond angle criteria for hydrogen bonds (Szyperski et al., 1992). Third, this hydrogen bond has also been observed in the X-ray crystal structure of the hirudin–thrombin complex (Rydel et al., 1990, 1991), where the carboxylate of Glu 17 forms both an intermolecular salt bridge to the guanidinium moiety of Arg 173 in thrombin and an intramolecular hydrogen bond with the amide proton of Cys 39.

The hydrogen bond involving the backbone amide proton of Gly 25 and the carboxylate of Glu 43 connects the type II turn comprising residues 23–26 with the polypeptide segment 41–45, which links the second strand of the $\beta$-sheet 27–31 and 36–40 with the short stretch of polyproline helix II containing

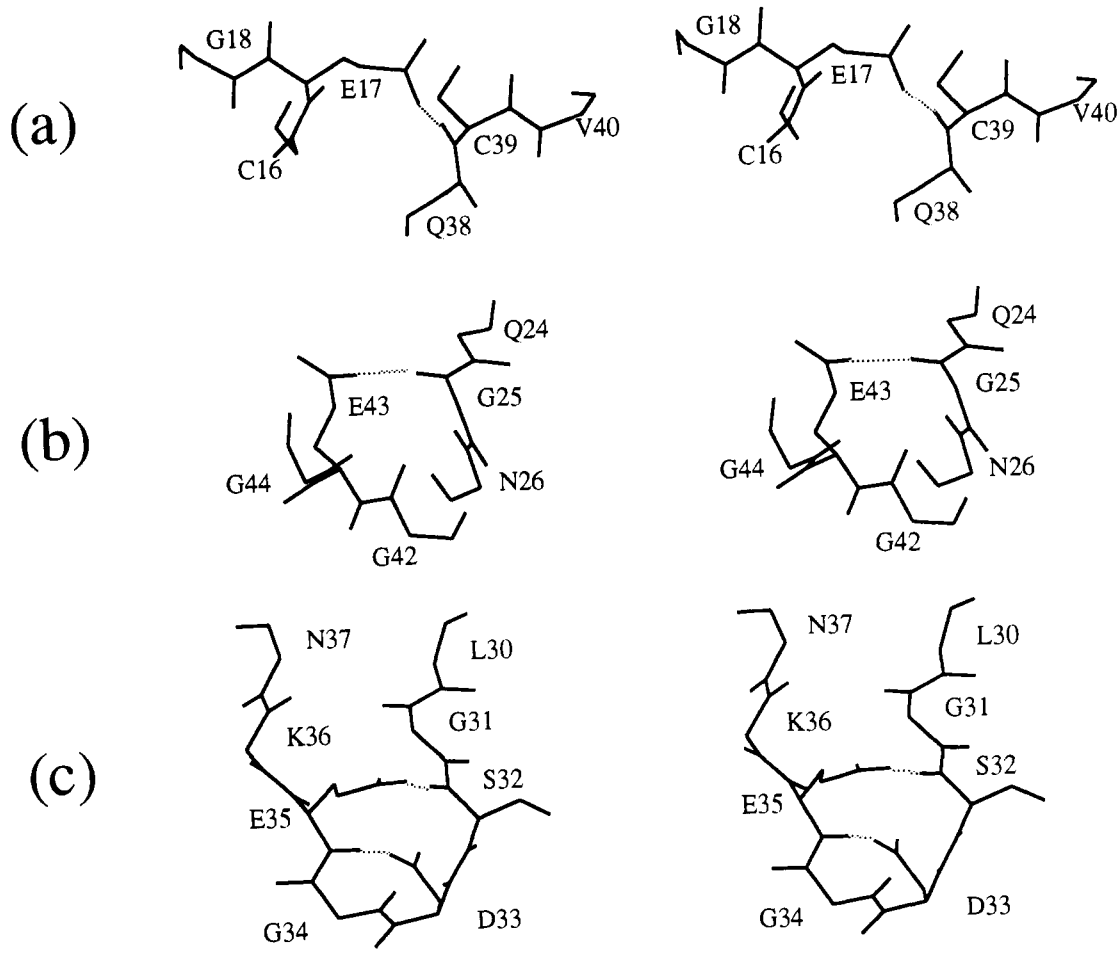


FIGURE 8: Stereo pictures of the hydrogen bonds (indicated as dotted lines) involving backbone amide protons and side-chain carboxylates in hirudin(1–51) as obtained from a structure calculation using hydrogen bond constraints (Williamson et al., 1985) in addition to the input of NOE upper distance constraints and dihedral angle constraints from spin-spin coupling constants (see text): (a) Cys 39 NH–Glu 17 $\delta COO^-$; (b) Gly 25 NH–Glu 43 $\delta COO^-$; (c) Ser 32 NH–Glu 35 $\delta COO^-$ and Glu 35 NH–Asp 33 $\gamma COO^-$. The drawings were generated with the program XAM (Xia, 1992).

residues 46–48. While the type II turn is locally and globally well defined in the NMR solution structure of hirudin(1–51), the polypeptide segment 41–45 exhibits local conformational disorder (see Figures 2 and 3 in Szyperski et al. (1992)). The downfield shift of the backbone amide proton of Gly 25 is about 3 times smaller when compared with that of Cys 39, and the hydrogen bond was identified neither in the NMR solution structure of hirudin(1–51) (Szyperski et al., 1992) nor in the X-ray crystal structure of the hirudin–thrombin complex (Rydel et al., 1990, 1991), suggesting that the increased entropic expense for the fixation of the polypeptide segment 41–45 results in a reduced population of the hydrogen bond.

From entropy considerations, relatively low populations are expected also for the hydrogen bonds involving the backbone amide protons of Glu 35 and Ser 32 and the carboxylates of Glu 35 and Asp 33. All these residues are located in the flexibly disordered loop comprising residues 31–36. The amide proton downfield shifts are again smaller when compared with that of Gly 25, and none of these hydrogen bonds was identified in the NMR solution structure of hirudin(1–51) (Szyperski et al., 1992). In addition, the polypeptide segment 32–35 of hirudin in the X-ray crystal structure of the hirudin–thrombin complex exhibits no electron density, implicating structural disorder for these residues also in the crystals (Rydel et al., 1990, 1991).

It was previously suggested that the extent of the amide proton downfield titration shift might be used to estimate the population of the hydrogen bond with the carboxylate group (Bundi & Wüthrich, 1979). The resulting numbers must be used with care, however, since the shift will depend on the hydrogen bond geometry as well as on the population. If we assume that the shift of 1.45 ppm for Cys 39 reflects a population of 70–90%, we obtain the following approximate populations for the other interactions: Gly 25 NH–Glu 43 $\delta COO^-$, ~30%; Ser 32 NH–Glu 35 $\delta COO^-$, ~20%; Glu 35 NH–Asp 33 $\gamma COO^-$, ~10%; Glu 35 NH–Glu 35 $\delta COO^-$, ~10%. The equilibrium populations of the hydrogen bonds appear to depend critically on the local mobilities of the interacting groups in the protein structure, since the enthalpic gain from hydrogen bond formation on the protein surface can apparently compensate the entropic expense for the fixation of a single side chain but not the entropic expense for immobilization of more extended polypeptide segments. This conclusion coincides with the observations of Dao-Pin et al. (1991) that engineered surface salt bridges between charged side chains do not contribute significantly to the protein stability, because the entropic cost of localizing a pair of solvent-exposed charged groups on the surface largely offsets the interaction energy expected from the formation of a defined salt bridge. Consistent with these general conclusions, no significantly populated hydrogen-bonding interactions between carboxylates and side-chain amide groups are implicated by the data on hirudin, suggesting that the enthalpic contribution of such an interaction is too small to act as a driving force for the formation of a defined hydrogen bond. It can thus be

predicted that significantly populated hydrogen-bonding interactions between side-chain carboxylates and side-chain amide groups are scarce on protein surfaces in solution, whereas hydrogen bonds between side-chain carboxylates and backbone amide protons can be expected whenever the backbone atoms of the interacting residues are in a sterically favorable, fixed conformation. The hydrogen bond between the carboxylate of Glu 17 and the amide proton of Cys 39 appears to establish a paradigm for this latter situation which also indicates that in earlier work the populations of such hydrogen bonds were overestimated about 2-fold based on the extent of $\Delta\delta(H^N)$ (Bundi & Wüthrich, 1979).

Clearly, considerable effort was needed both for the preparation of NMR quantities of the mutant proteins and for the recording of the NMR data (Table 2). However, there is no alternative approach for obtaining the results described here. Although screening the NMR solution structure with the common bond length and bond angle criteria resulted in the identification of the secondary structure hydrogen bonds, the Cys 39 NH–Glu 17 $\delta COO^-$ bond was the only surface interaction thus found. The most important information is undoubtedly on the dynamic nature of the surface interactions. In all instances the spectra of the hydrogen-bonded and nonbonded forms of both the donor and acceptor groups are averaged on the chemical shift time scale (Wüthrich, 1986), which shows that the lifetimes of the hydrogen bonds are in the millisecond range or shorter. In particular, the amide proton of Glu 35 must exchange between two different hydrogen–bonded states with Glu 35 $\delta COO^-$ and Asp 33 $\gamma COO^-$ and additional, nonbonded states, which presents a novel, dynamic picture of a "bifurcated" hydrogen bond. Although no detailed data are presently available, the sizeable downfield titration shifts of amide protons in the flexible C-terminal tail from residues 53–65 (Figure 2) would be compatible with the formation of possibly a multitude of rapidly interchanging hydrogen bonds with the numerous carboxylate groups in this polypeptide segment.

Finally, it may be added that the protocol of NMR pH titrations and comparison of single-residue mutations used here for desulfatohirudin and previously with model peptides (Bundi & Wüthrich, 1979) is generally applicable for studies of the surface structure in polypeptides and proteins that do not undergo global conformation changes over the pH range 2–6.

## ACKNOWLEDGMENT

We thank Dr. P. Güntert for help with the structure calculations, Dr. M. Billeter for helpful discussions, Dr. H. Grossenbacher (Ciba-Geigy AG, Basel) for a generous gift of desulfatohirudin, and Mrs. E. Huber for the careful processing of the manuscript.

## REFERENCES

Anil-Kumar, Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.

Badgy, D., Barabas, E., Graf, L., Ellebaek, T., & Magnusson, S. (1976) *Methods Enzymol. 45*, 669–678.

Braun, P. J., Dennis, S., Hofsteenge, J., & Stone, S. R. (1988) *Biochemistry 27*, 6517–6522.

Bundi, A., & Wüthrich, K. (1977) *FEBS Lett. 77*, 11–14.

Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 299–311.

Close, P., Bichler, J., Kerry, R., et al. (1994) *Coronary Artery Disease* (in press).

Dao-Pin, S., Sauer, U., Nicholson, H., & Matthews, B. W. (1991) *Biochemistry 30*, 7142–7153.

DeMarco, A., & Wüthrich, K. (1976) *J. Magn. Reson. 24*, 201–204.

Dennis, S., Wallace, A., Hofsteenge, J., & Stone, S. R. (1990) *Eur. J. Biochem. 188*, 61–66.

Dodt, J., Müller, H. P., Seemüller, V., & Chang, J.-Y. (1984) *FEBS Lett. 165*, 180–184.

Ebina, S., & Wüthrich, K. (1984) *J. Mol. Biol. 179*, 283–288.

Folkers, P. J. M, Clore, G. M., Driscoll, P. C., Dodt, J., Köhler, S., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2601–2617.

Griesinger, C., Otting, G., Wüthrich, K., & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870-7872.

Grossenbacher, H., Auden, J. A. L., Bill, K., Liersch, M., & Maerki, W. E. (1987) 11th Congress on Thrombosis and Haemostasis, Brussels, July 11, 1987, Abstract 34.

Güntert, P., Braun, W., & Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517–530.

Haruyama, H., & Wüthrich, K. (1989) *Biochemistry 28*, 4301–4312.

Haruyama, H., Qian, Y.-Q., & Wüthrich, K. (1989) *Biochemistry 28*, 4312–4317.

Johnson, P. H., Sze, P., Winant, R., Payne, P. W., & Lazar, J. B. (1989) *Semin. Thromb. Hemostasis 15*, 302–315.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.

Langone, J. J., Ed. (1989) *Methods in Enzymology*, Vol. 178, Academic Press, San Diego.

Lent, C. (1986) *Nature 323*, 494.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Marion, D., & Bax, A. (1988) *J. Magn. Reson. 80*, 528–533.

Marion, D., Ikura, K., & Bax, A. (1989) *J. Magn. Reson. 84*, 425–430.

Märki, W. E., & Wallis, R. B. (1990) *Thromb. Haemostasis 64*, 344–348.

Meyhack, B., Heim, J., Rink, H., Zimmermann, W., & Maerki, W. E. (1987) 11th Congress on Thrombosis and Haemostasis, Brussels, July 11, 1987, Abstract 33.

O'Connell, J. F., Bougis, P. E., & Wüthrich, K. (1993) *Eur. J. Biochem. 213*, 891–900.

Rydel, T. J., Ravichandran, K. G., Tulinsky, A., Bode, W., Huber, R., Roitsch, C., & Fenton, J. W., II (1990) *Science 249*, 277–280.

Rydel, T. J., Tulinsky, A., Bode, W., & Huber, R. (1991) *J. Mol. Biol. 221*, 583–601.

Sasisekharan, V. (1959) *Acta Crystallogr. 12*, 897–903.

Scharf, M., Engels, M., & Tripier, D. (1989) *FEBS Lett. 255*, 105–110.

Steinmetz, W. E., Bougis, P. E., Rochat, H., Redwine, O. D., Braun, W., & Wüthrich, K. (1988) *Eur. J. Biochem. 172*, 101–116.

Stone, S. R., & Hofsteenge, J. (1986) *Biochemistry 25*, 4622–4624.

Szyperski, T., Güntert, P., Stone. S. R., & Wüthrich, K. (1992) *J. Mol. Biol. 228*, 1193–1205.

Williamson, M. P., Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 182*, 295–315.

Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, Wiley, New York.

Xia, T. (1992) Ph.D. Thesis 9831, ETH Zürich, Switzerland.