**Supporting Information**

# 3D protein structure determination using pseudocontact shifts of backbone amide protons generated by double-histidine $Co^{2+}$-binding motifs at multiple sites

*Alireza Bahramzadeh, Thomas Huber*, and Gottfried Otting**

Research School of Chemistry, Australian National University, Canberra, ACT 2601 (Australia)

**Table S1.** $\Delta\chi$ tensor parameters of ERp29-C fitted to the structure 2M66[1] using the PCSs measured with dHis-$Co^{2+}$ motifs.[a]

| | | Mu1 | Mu2 | Mu3 | Mu4 |
|---|---|---|---|---|---|
| $\Delta\chi_{ax}$[b] | | -2.3 | -1.8 | -3.1 | 3.1 |
| $\Delta\chi_{rh}$[b] | | -0.7 | -1.1 | -1.0 | 1.0 |
| Metal coordinates[c] | x | 4.957 | -3.581 | 13.597 | 23.499 |
| | y | 17.411 | 8.383 | 16.046 | 13.816 |
| | z | 11.492 | -10.399 | 14.731 | -4.848 |
| Euler angles (degrees)[c,d] | $\alpha$ | 155 | 77 | 158 | 75 |
| | $\beta$ | 122 | 133 | 96 | 156 |
| | $\gamma$ | 134 | 55 | 24 | 128 |
| SD ($ppm^2$)[e] | | 0.031 | 0.012 | 0.039 | 0.043 |
| RMSD (ppm)[f] | | 0.020 | 0.013 | 0.021 | 0.027 |
| *Q*-factor[g] | | 0.24 | 0.21 | 0.23 | 0.27 |

[a] The program Numbat[2] was used to fit the $\Delta\chi$ tensors. The fit used all residues of all 11 conformers of 2M66 to determine the $\Delta\chi$ tensor, which yielded the lowest squared deviation between back-calculated and experimental PCSs. The fits were performed without restricting the position of the metal ion. The experimental PCSs are reported in Table S2.

[b] In units of $10^{-32}$ $m^3$.

[c] Relative to the coordinates reported in the PDB file 2M66.

[d] Expressed in the 'ZYZ' convention.

[e] Sum of the square deviation of back-calculated and experimental PCSs.

[f] Root mean square deviation of back-calculated and experimental PCSs.

[g] Calculated as $\mathrm{Sqrt}[\Sigma(PCS_{exp} - PCS_{calc})^2/\Sigma(PCS_{exp})^2]$, where $PCS_{exp}$ and $PCS_{calc}$ are the experimental and back-calculated PCSs, respectively.

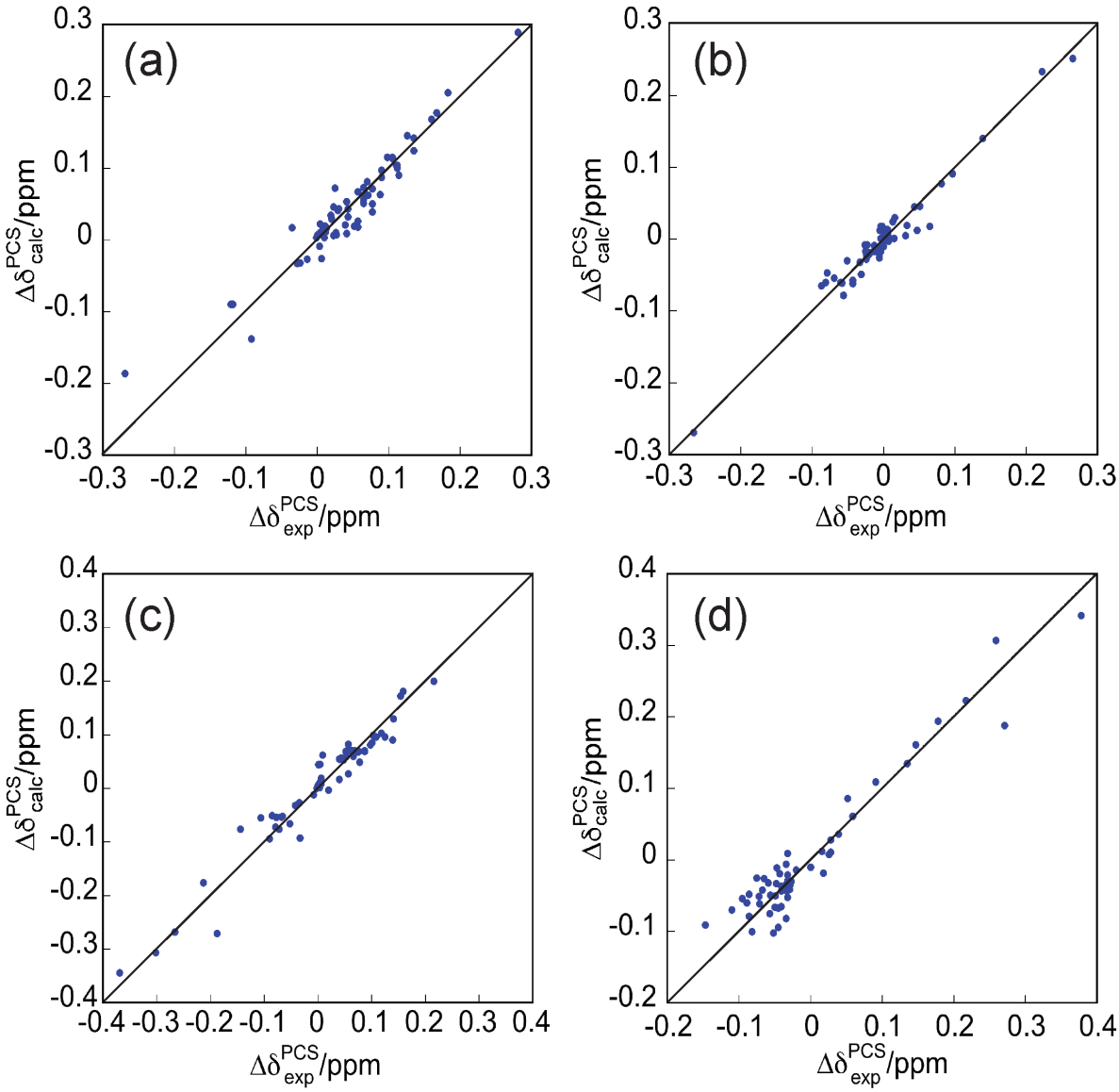


**Figure S1.** Correlation between back-calculated and experimental PCSs for the four different mutants of ERp29-C (Mu1 – Mu4), each of which contains a dHis-$Co^{2+}$ motif at a different site. The $\Delta\chi$ tensors of Table S1 were used for the back-calculation of the PCSs.

**Table S2.** Experimental PCSs of the backbone amide protons of ERp29-C with a dHis-$Co^{2+}$ motif at four different sites.[a]

| Total number of PCSs | 73 | 69 | 85 | 58 |
|---|---|---|---|---|
| Residue | PCSs of Mu1 | PCSs of Mu2 | PCSs of Mu3 | PCSs of Mu4 |
| G156 | | | 0.001 | |
| C157 | | | 0.009 | |
| L158 | | -0.087 | 0.052 | |
| A160 | | -0.031 | | |
| Y161 | | -0.043 | 0.125 | |
| D162 | | -0.056 | 0.020 | |
| A163 | | -0.058 | -0.033 | -0.032 |
| L164 | | -0.079 | | |
| A165 | | -0.06 | -0.090 | |
| G166 | | -0.081 | -0.266 | -0.034 |
| Q167 | | -0.033 | | |
| F168 | | -0.006 | -0.369 | 0.026 |
| I169 | | | -0.213 | 0.002 |
| E170 | | | -0.302 | 0.018 |
| A171 | | 0.007 | -0.188 | 0 |
| S172 | 0.281 | 0.004 | 0.004 | -0.043 |
| S173 | 0.183 | 0.006 | | -0.107 |
| R174 | 0.090 | 0.031 | | -0.046 |
| E175 | 0.077 | 0.003 | | -0.050 |
| A176 | 0.112 | 0.003 | | |
| R177 | | 0.015 | | |
| Q178 | 0.114 | 0.004 | | 0.059 |
| A179 | | 0.002 | | |
| I180 | | -0.002 | | |
| L181 | 0.105 | -0.013 | | |
| K182 | -0.028 | -0.001 | | |
| Q183 | | 0 | | |
| G184 | -0.024 | -0.008 | | 0.217 |
| Q185 | -0.001 | -0.009 | | 0.378 |
| D186 | -0.118 | -0.006 | | 0.259 |
| G187 | -0.121 | -0.01 | | 0.178 |
| L188 | 0.003 | -0.01 | 0.154 | |
| S189 | -0.014 | -0.009 | 0.139 | 0.147 |
| G190 | 0.006 | -0.01 | 0.216 | 0.091 |
| V191 | 0.057 | -0.011 | 0.159 | 0.034 |
| K192 | 0.088 | -0.024 | 0.141 | 0.039 |
| E193 | 0.043 | -0.012 | 0.067 | 0.028 |
| T194 | 0.043 | -0.013 | 0.044 | 0.016 |

| | | | | |
|---|---|---|---|---|
| D195 | 0.071 | -0.020 | 0.069 | 0.028 |
| K196 | 0.067 | -0.020 | 0.062 | |
| K197 | 0.057 | -0.025 | 0.048 | -0.047 |
| W198 | 0.090 | -0.024 | 0.053 | -0.075 |
| A199 | 0.135 | -0.024 | 0.057 | -0.109 |
| S200 | 0.111 | -0.008 | 0.076 | -0.020 |
| Q201 | | -0.004 | 0.040 | -0.045 |
| Y202 | 0.160 | -0.051 | 0.064 | |
| L203 | 0.167 | -0.022 | 0.109 | -0.059 |
| K204 | | -0.023 | 0.087 | |
| I205 | 0.135 | -0.026 | 0.053 | |
| M206 | | -0.013 | 0.098 | |
| G207 | 0.126 | 0.007 | 0.103 | |
| K208 | 0.112 | -0.005 | 0.087 | |
| I209 | 0.106 | | 0.075 | |
| L210 | 0.098 | | 0.118 | |
| D211 | 0.065 | 0.047 | 0.101 | |
| Q212 | 0.065 | 0.065 | 0.066 | |
| G213 | 0.065 | -0.004 | 0.078 | |
| E214 | 0.052 | -0.002 | 0.057 | |
| D215 | 0.041 | | 0.040 | 0.013 |
| F216 | 0.057 | 0.043 | 0.006 | |
| A218 | 0.023 | | 0.001 | -0.052 |
| S219 | 0.027 | | 0.003 | -0.147 |
| E220 | 0.039 | | 0.005 | |
| L221 | 0.013 | | 0.001 | -0.05 |
| A222 | 0.010 | | -0.002 | -0.11 |
| R223 | 0.026 | | 0 | -0.057 |
| I224 | 0.016 | | 0 | -0.034 |
| S225 | 0.008 | | -0.001 | -0.089 |
| K226 | 0.008 | | 0 | -0.056 |
| L227 | 0.007 | | 0.001 | -0.095 |
| I228 | 0.012 | 0.265 | 0 | -0.086 |
| E229 | 0 | 0.222 | 0 | -0.041 |
| N230 | 0.002 | 0.139 | 0.003 | -0.028 |
| K231 | 0.001 | 0.097 | 0 | -0.032 |
| M232 | 0.004 | 0.081 | 0 | -0.048 |
| S233 | 0.012 | 0.051 | 0.003 | -0.037 |
| E234 | 0.011 | 0.033 | 0.002 | -0.027 |
| G235 | 0.011 | 0.006 | 0.001 | -0.034 |
| K236 | 0.011 | 0.013 | 0.003 | -0.039 |
| K237 | 0.008 | 0.016 | 0.004 | -0.033 |
| E238 | 0.004 | 0 | 0.004 | -0.067 |
| E239 | 0.020 | -0.004 | 0.006 | -0.032 |

| | | | | |
|---|---|---|---|---|
| L240 | 0.020 | -0.002 | 0.006 | -0.008 |
| Q241 | 0.020 | -0.006 | 0.002 | 0.019 |
| R242 | 0.029 | -0.043 | 0.005 | |
| S243 | 0.041 | -0.069 | 0.004 | |
| L244 | 0.030 | | 0.001 | |
| N245 | 0.023 | | 0.003 | |
| I246 | 0.070 | | 0.001 | -0.041 |
| L247 | 0.025 | -0.266 | 0 | |
| T248 | 0.019 | | -0.008 | |
| A249 | 0.077 | | -0.034 | |
| F250 | 0.077 | | -0.038 | -0.045 |
| R251 | | | -0.042 | |
| K252 | -0.092 | | -0.068 | -0.049 |
| K253 | | | -0.144 | |
| G254 | -0.269 | | -0.106 | -0.072 |
| A255 | | | -0.085 | -0.029 |
| E256 | | | -0.079 | |
| K257 | | | -0.072 | |
| E258 | | | -0.066 | |
| E259 | | | -0.077 | -0.032 |
| L260 | | | -0.052 | |

[a] PCSs were measured in ppm by subtracting the chemical shift of samples with $Co^{2+}$ from samples with $Zn^{2+}$ observed in [$^{15}$N,$^{1}$H]-HSQC spectra of uniformly $^{15}$N/$^{13}$C-labeled samples at 25 $^{o}$C in 20 mM MES buffer, pH 6.5, containing 50 mM NaCl buffer.

**Table S3.** Δχ tensor parameters and cobalt ion coordinates relative to the first conformer of 2M66 and the individual conformers of the Top10 ensemble.[a]

| Mutant | | 2M66 | top1 | top2 | top3 | top4 | top5 | top6 | top7 | top8 | top9 | top10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mu1** | | | | | | | | | | | | |
| $\Delta\chi_{ax}$[b] | | -2.63 | -2.79 | -2.52 | -2.66 | -2.70 | -2.61 | -2.81 | -2.70 | -3.02 | -3.16 | -3.05 |
| $\Delta\chi_{rh}$[b] | | -0.38 | -0.34 | -0.22 | -0.54 | -0.51 | -0.57 | -0.35 | -0.62 | -0.77 | -0.77 | -0.68 |
| Metal coordinates | x | 5.80 | 4.58 | 4.17 | 6.62 | 4.01 | 4.23 | 3.63 | 3.75 | 3.94 | 4.30 | 4.62 |
| | y | 15.41 | 16.78 | 17.62 | 16.86 | 16.46 | 17.79 | 17.20 | 17.16 | 16.74 | 16.68 | 17.09 |
| | z | 14.05 | 13.39 | 13.14 | 13.49 | 13.55 | 13.04 | 13.18 | 12.81 | 13.30 | 13.58 | 13.47 |
| Euler angles (degrees) | α | 165 | 163 | 162 | 159 | 159 | 148 | 162 | 157 | 157 | 159 | 158 |
| | β | 112 | 115 | 116 | 108 | 115 | 118 | 117 | 120 | 119 | 116 | 116 |
| | γ | 145 | 150 | 140 | 155 | 135 | 146 | 149 | 143 | 149 | 156 | 154 |
| SD | | 0.05 | 0.02 | 0.04 | 0.04 | 0.03 | 0.07 | 0.03 | 0.04 | 0.04 | 0.03 | 0.04 |
| *Q*-factor[c] | | 0.32 | 0.18 | 0.28 | 0.26 | 0.25 | 0.35 | 0.23 | 0.28 | 0.29 | 0.25 | 0.27 |
| **Mu2** | | | | | | | | | | | | |
| $\Delta\chi_{ax}$[b] | | -2.43 | -2.20 | -2.66 | -2.58 | -2.29 | -2.40 | -2.61 | -2.45 | -2.52 | -2.37 | -2.32 |
| $\Delta\chi_{rh}$[b] | | -1.54 | -1.21 | -1.54 | -1.15 | -1.08 | -1.45 | -1.68 | -1.57 | -1.54 | -1.08 | -1.45 |
| Metal coordinates | x | -2.66 | -1.38 | -1.07 | -0.55 | -0.91 | -4.28 | -1.78 | -2.06 | -1.60 | 0.35 | -3.06 |
| | y | 19.92 | 19.65 | 20.21 | 20.26 | 18.99 | 18.96 | 20.48 | 19.62 | 19.57 | 20.89 | 19.60 |
| | z | -11.84 | -13.01 | -14.34 | -14.39 | -15.18 | -10.69 | -12.66 | -12.15 | -12.97 | -15.06 | -10.91 |
| Euler angles (degrees) | α | 66 | 67 | 60 | 59 | 52 | 80 | 62 | 63 | 60 | 58 | 75 |
| | β | 130 | 142 | 142 | 141 | 152 | 130 | 138 | 136 | 136 | 151 | 135 |
| | γ | 51 | 50 | 45 | 39 | 38 | 62 | 54 | 49 | 45 | 54 | 65 |
| SD | | 0.01 | 0.03 | 0.03 | 0.02 | 0.03 | 0.02 | 0.04 | 0.03 | 0.02 | 0.06 | 0.04 |
| *Q*-factor[c] | | 0.22 | 0.29 | 0.31 | 0.25 | 0.32 | 0.27 | 0.38 | 0.30 | 0.24 | 0.45 | 0.35 |
| **Mu3** | | | | | | | | | | | | |
| $\Delta\chi_{ax}$[b] | | -3.94 | -4.10 | -3.51 | -2.64 | -3.87 | -3.64 | -3.95 | -3.23 | -3.67 | -3.39 | -3.97 |
| $\Delta\chi_{rh}$[b] | | -2.08 | -1.78 | -1.22 | -0.77 | -1.70 | -1.43 | -1.64 | -1.17 | -1.44 | -1.34 | -1.81 |
| Metal coordinates | x | 17.78 | 14.78 | 14.22 | 13.74 | 14.47 | 14.17 | 14.49 | 13.78 | 14.02 | 14.35 | 15.01 |
| | y | 16.63 | 15.96 | 15.96 | 16.34 | 15.43 | 16.08 | 15.77 | 15.75 | 16.20 | 16.11 | 15.93 |
| | z | 14.35 | 14.75 | 14.96 | 14.02 | 14.84 | 14.64 | 14.73 | 14.80 | 14.67 | 14.52 | 14.57 |
| Euler angles (degrees) | α | 154 | 150 | 152 | 157 | 153 | 151 | 152 | 154 | 154 | 155 | 150 |
| | β | 84 | 94 | 98 | 96 | 92 | 95 | 93 | 99 | 97 | 92 | 90 |
| | γ | 44 | 28 | 28 | 22 | 26 | 27 | 27 | 24 | 27 | 28 | 32 |
| SD | | 0.11 | 0.03 | 0.07 | 0.05 | 0.06 | 0.05 | 0.04 | 0.07 | 0.03 | 0.06 | 0.06 |
| *Q*-factor[c] | | 0.37 | 0.19 | 0.29 | 0.25 | 0.27 | 0.25 | 0.24 | 0.30 | 0.20 | 0.27 | 0.29 |

| **Mu4** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\Delta\chi_{ax}$[b] | | 1.84 | 2.65 | 2.26 | 2.12 | 2.21 | 2.17 | 1.99 | 2.68 | 2.30 | 1.71 | 2.04 |
| $\Delta\chi_{rh}$[b] | | 1.12 | 1.39 | 1.19 | 1.08 | 1.20 | 1.23 | 1.15 | 1.48 | 1.14 | 0.95 | 1.18 |
| Metal coordinates | x | 20.47 | 21.00 | 19.81 | 20.21 | 20.79 | 20.54 | 20.50 | 20.28 | 21.07 | 20.78 | 20.47 |
| | y | 15.95 | 16.00 | 15.90 | 15.58 | 16.14 | 16.06 | 15.70 | 16.41 | 14.94 | 15.98 | 16.98 |
| | z | -2.56 | -5.00 | -4.98 | -4.54 | -3.75 | -3.82 | -3.44 | -5.45 | -4.32 | -2.16 | -2.80 |
| Euler angles (degrees) | α | 92 | 95 | 96 | 91 | 90 | 89 | 88 | 101 | 86 | 80 | 91 |
| | β | 139 | 145 | 147 | 145 | 141 | 141 | 142 | 144 | 149 | 142 | 134 |
| | γ | 145 | 137 | 140 | 142 | 136 | 141 | 137 | 140 | 134 | 134 | 143 |
| SD | | 0.07 | 0.07 | 0.06 | 0.06 | 0.06 | 0.07 | 0.10 | 0.09 | 0.11 | 0.06 | 0.10 |
| *Q*-factor[c] | | 0.35 | 0.35 | 0.33 | 0.33 | 0.31 | 0.35 | 0.41 | 0.40 | 0.44 | 0.32 | 0.42 |

[a] The metal coordinates were determined by modeling the dHis-$Co^{2+}$ motif using a PyParaTools[3] module, which finds the metal coordinates with minimal RMSD with regard to the expected carbon-metal distances of the dHis-$Co^{2+}$ motif (Figure 3b). The remaining $\Delta\chi$ tensor parameters were determined using the program Numbat[2] with the metal position fixed to these coordinates.

[b] In units of $10^{-32}$ $m^3$.

[c] Calculated as the root-mean-square (RMS) of the differences between experimental and back-calculated PCSs divided by the RMS of the experimental PCSs.

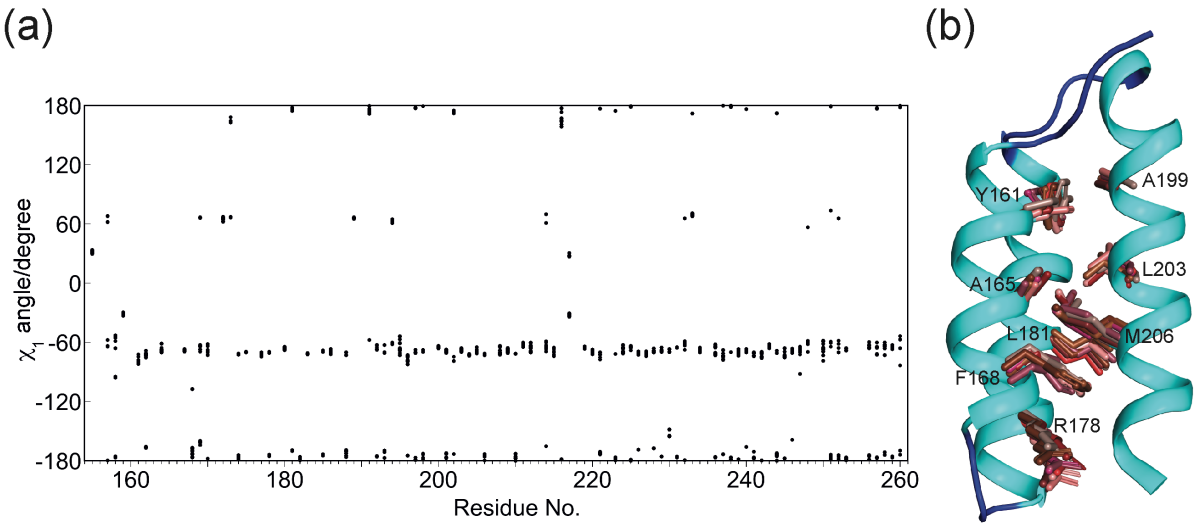


**Figure S2.** Side chain conformations determined by the GPS-Rosetta calculations. $\chi_1$ angles of all residues in the Top10 conformers. (b) Illustration of precision of side chain conformations of buried residues from helices 5–7 in the Top10 conformers (plotted in shades of red) following $C^{\alpha}$ atom superimposition of the conformers.

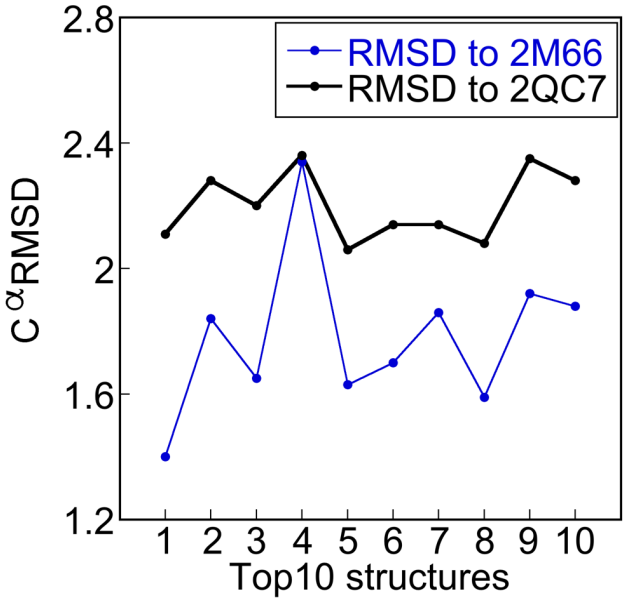


**Figure S3.** Comparison of the Top10 conformers to the first conformer of the structure 2M66[1] and the crystal structure 2QC7.[4] The figure reports the $C^{\alpha}$ RMSD values determined by PyMOL[5] following superimposition for minimal RMSD. The RMSD values are consistently smaller with respect to 2M66 than 2QC7, with the conformer top1 showing the lowest RMSD to 2M66 (95 residues aligned, comprising residues 156-251 of rat ERp29-C; the rat protein contains one residue less than the human protein (deletion of K230)).

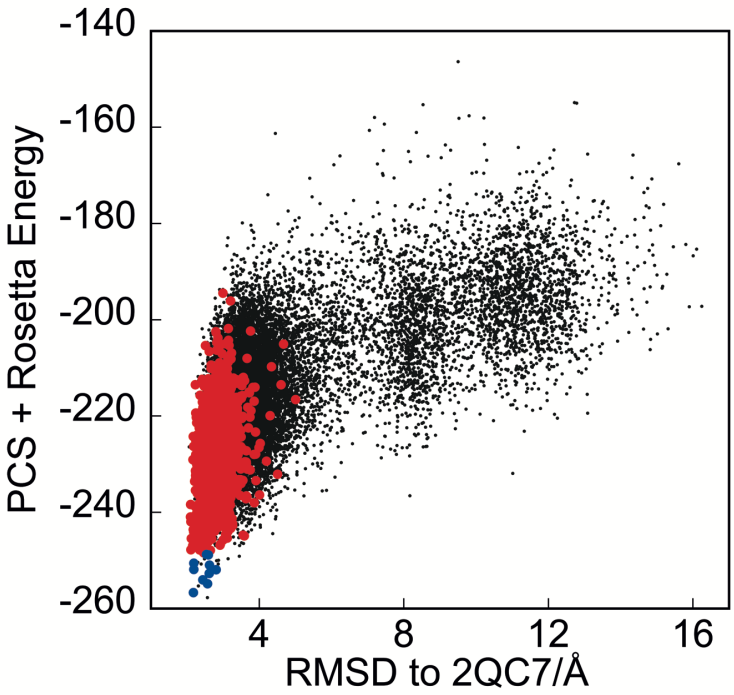


**Figure S4.** Combined PCS + Rosetta energy scores of the 20,000 GPS-Rosetta structural models (black points) calculated in the present work plotted versus the $C^{\alpha}$ RMSD to the crystal structure of human ERp29 (PDB ID: 2QC7). The alignment included all residues of the C-terminal domain in the crystal structure (155–256). Black points mark the result for all 20,000 structures obtained from GPS-Rosetta. Blue and red points mark the same structures as in Figure 3a, i.e. structures fulfilling the PCS restraints as well as the distance restraints specified in Figure 3b.

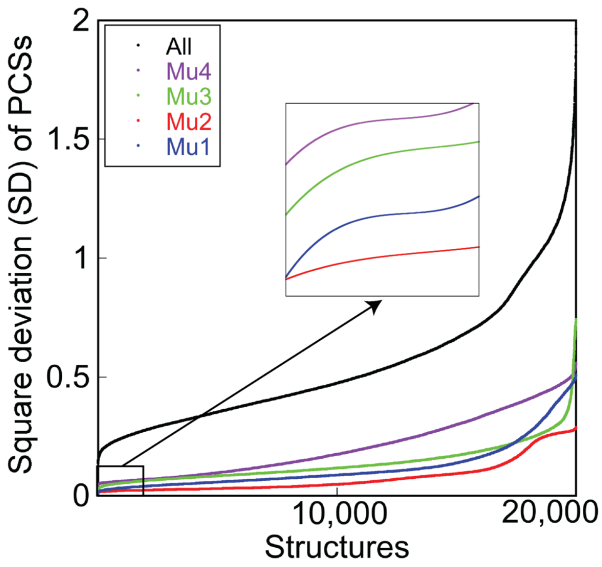


**Figure S5.** Performance of the PCS data set from each dHis-$Co^{2+}$ motif in the GPS-Rosetta calculations. The GPS-Rosetta models calculated were sorted by their square deviations (SD) between experimental and back-calculated PCSs. The coordinated descent of SD values observed for the four mutants indicates that all four PCS datasets contributed to the overall convergence of the GPS-Rosetta calculations.

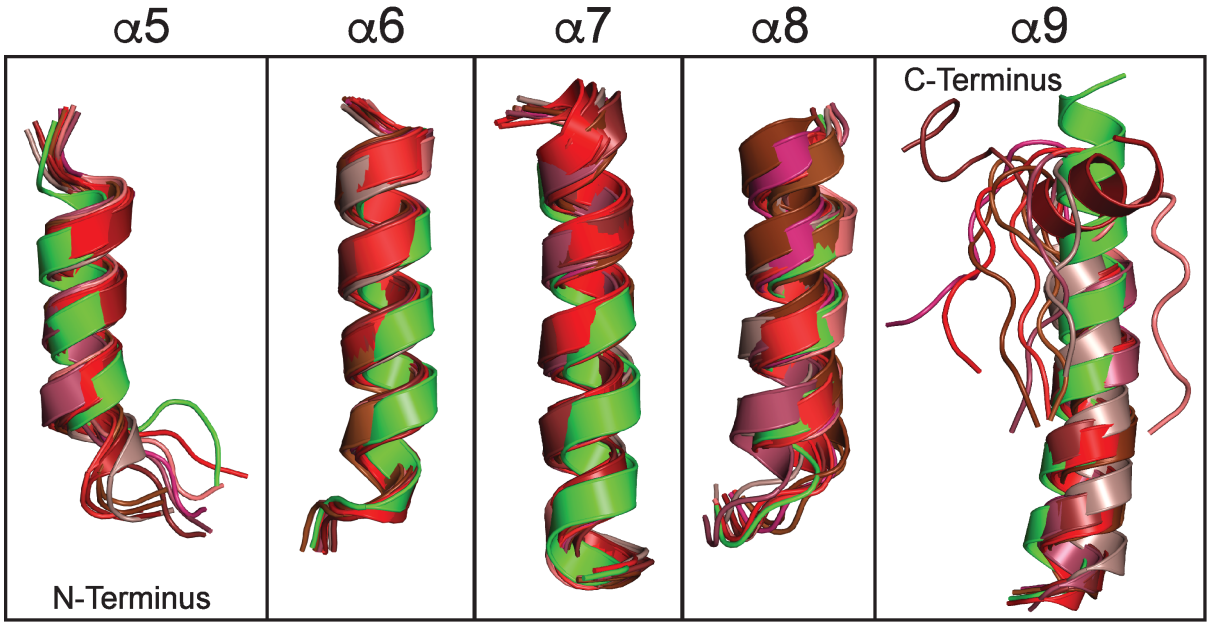


**Figure S6.** Comparison of the individual α-helices of the Top10 conformers (plotted in shades of red) with the first conformer of 2M66 (green) following global superimposition of the structures. Helices 5-7 align more closely than helices 8 and 9.

## References

(1) Yagi, H., Pilla, K. B., Maleckis, A., Graham, B., Huber, T., and Otting, G. (2013) Three-dimensional protein fold determination from backbone amide pseudocontact shifts generated by lanthanide tags at multiple sites. *Structure 21*, 883–890.

(2) Schmitz, C., Stanton-Cook, M. J., Su, X.-C., Otting, G., and Huber, T. (2008) Numbat: an interactive software tool for fitting Δχ-tensors to molecular coordinates using pseudocontact shifts. *J. Biomol. NMR 41*, 179–189.

(3) Stanton-Cook, M. J., Su, X.-C., Otting, G., and Huber, T. (2010) PyParaTools. http://comp-bio.anu.edu.au/mscook/PPT/, accessed April 25, 2019.

(4) Barak, N. N., Neumann, P., Sevvana, M., Schutkowski, M., Naumann, K., Malešević, M., Reichardt, H., Fischer, G., Stubbs, M. T., and Ferrari, D. M. (2009) Crystal structure and functional analysis of the protein disulfide isomerase-related protein ERp29. *J. Mol. Biol. 385*, 1630–1642.

(5) The PyMOL Molecular Graphics system, Version 2.0 Schrödinger, LLC.