THE SOLUTION STRUCTURE OF DNA-FREE PAX-8 PAIRED BOX DOMAIN ACCOUNTS FOR REDOX REGULATION OF TRANSCRIPTIONAL ACTIVITY IN PAX PROTEIN FAMILY

Luca Codutti[1], Hugo van Ingen[2], Carlo Vascotto[1], Federico Fogolari[1,3], Alessandra Corazza[1,3], Gianluca Tell[1], Franco Quadrifoglio[1], Paolo Viglino[1,3], Rolf Boelens[2] and Gennaro Esposito[1,3]

From Dipartimento di Scienze e Tecnologie Biomediche, Università degli Studi di Udine p.le Kolbe 4, 33100 Udine, Italy[1]; Bijvoet Center for Biomolecular Research Padualaan 8, 3584 CH Utrecht, The Netherlands[2]; INBB (Istituto Nazionale Biostrutture e Biosistemi) Viale Medaglie d'Oro 35, 00100 Roma , Italy [3]

Running head: *Pax-8 Paired Box Domain solution structure*

Address correspondence to: Prof. G. Esposito, Dipartimento di Scienze e Tecnologie Biomediche, Università degli Studi di Udine p.le Kolbe 4, 33100 Udine, Italy. Fax: ++390432494301. E-mail: gesposito@mail.dstb.uniud.it

SUPPLEMENTAL DATA

PROTEIN EXPRESSION AND PURIFICATION

E. coli BL21 (DE3) pLysS competent cells (Stratagene) are transformed through heat-shock then incubated overnight, using ampicilline and chloramphenicol as selective controls. Incubation in LB media + 1 % D-glucose follows, until reaching a 0.8 optical density (O.D.). Cells are then centrifuged and harvested and the resulting supernatant medium is replaced with a new medium E.coli OD2 CN (Silantes) containing 1 mM final IPTG.

Induction is performed for 8 hours at 310 K under shaking, and then cells are harvested by centrifugation. Bacteria are then lysed using a solution containing 20 mM Tris/HCl pH 8.00, 250 mM NaCl and 1% w/v Tween, and enriched with lysozyme, phenylmethanesulphonyl fluoride (PMSF), dithiothreitol (DTT) and protease inhibitors. Cells are kept in this solution at 310 K for 30 min, then sonicated and centrifuged.

Afterwards the supernatant fraction is purified by means of FPLC, using a chelating GE Healthcare HiTrap column, conditioned with a buffer composed of 20 mM Tris, 0.5 M NaCl, 0.5 mM DTT pH 7.4. A linear gradient from 0 to 100% of elution buffer (20 mM Tris, 0.5 M Imidazole, 0.5 M NaCl, 0.5 mM DTT pH 7.4) is then applied in 20 min, with a flux rate of 1 ml/min. The Prd domain purity was checked by means of ESI mass spectroscopy which confirmed the 100% incorporation of ^{13}C and ^{15}N, revealing a main m/z peak with mass equal to 18,403.7 Da (expected 18,402.8 Da).

NMR DATA ACQUISITION

All spectra were acquired, both at 500 and 750 MHz, with acquisition times in the range 50-120 ms to reach a suitable FID resolution in all dimensions. On average 2D ^{15}N-resolved two-dimensional spectra required 512 t increments of 1024 complex data points with 32-96 scans per transient whereas 2D ^{13}C-resolved experiments required a longer acquisition to reach a suitable resolution in the indirect dimension and typically resulted (at 750 MHz) in 600 increments, with 16 scans per transient. 3D triple frequency experiments acquired at 500 MHz were usually set up with 80x48x2048 complex data points with 64-196 scans, whereas 3D HSQC-TOCSY and NOESY experiments were set up with 264x44x2048 (^{15}N) and 160x80x2048 (^{13}C) points. Experiments acquired at 750 MHz generally required the acquisition of a higher number of indirect points, for example the 3D ^{13}C and ^{15}N resolved HSQC-NOESYs were respectively acquired with 460x80x1024 and 330x100x768 points with 8 scans.

Except for region-selective experiments, e.g. aromatic resonance NOEs, we used spectral widths around 7,000-10,000 Hz in the direct acquisition dimension (^{1}H), and around 2,500 Hz or 12,000-19,000 Hz in the ^{15}N or ^{13}C indirect dimensions, respectively. The nitrogen carrier frequency was set at 118.0 ppm to enable the aliasing of Arg N^{ε}, Gln $N^{\varepsilon 2}$ and Asn $N^{\delta 2}$ resonances into empty portions of the spectrum.

Multi-dimensional datasets were recorded in a phase-sensitive mode with quadrature detection by the States/TPPI [1, 2] or echo/antiecho method[3] in the indirectly detected dimensions. The ^{1}H carrier frequency was

set to the water resonance and broadband decoupling of the heteronuclei during acquisition was achieved using GARP or WALTZ-16 (1, 2), at decoupling field strengths corresponding to $\gamma_H B_1/2\pi$ of ~1,000 Hz (^{15}N) and ~ 2,500 Hz (^{13}C). All spectra were processed using Bruker TopSpin(R) 1.3. Original sizes were normally increased by linear prediction using the mirror-image or the backward-forward method, depending on the pulse sequence. Zero filling to double the data size followed by apodization with a shifted sine bell function was applied prior to Fourier transformation.

STRUCTURE CALCULATION DETAILS

Simulated annealing with 8,000 torsion angle dynamics time steps and 10,000 conjugate gradients minimization steps per conformer was performed. The 20 conformers with the lowest CYANA target function were subjected to refinement in implicit solvent (dielectric constant = 4 x r) using restrained energy minimization with the AMBER force field in the program DISCOVER (Accelrys) to improve local structure quality and electrostatics. Minimization was performed with 200 iterations using of steepest descent and 800 iterations using the conjugate gradients algorithm.

The force constants were chosen such that a restraint violation of 1 Å contributed 10 kcal/mol to the potential energy for the upper restraints and 2 kcal/mol for the lower restraints, while dihedral angle ranges violation of 1 degree, contributed 30 kcal/mol.

MD SIMULATION DETAILS

Protons were added to the crystal structure structure of human Pax-6 Prd Box domain (PDB code: 6PAX) by means of the pdb2gmx utility of the GROMACS simulation package (3). The protonated structure was then used to generate a topology and coordinate file through the psfgen utility of the NAMD simulation software (4). The structure was then used to compute the electrostatic potential around the molecule using the program UHBD (5). In order to neutralize the protein charge (+10) a chloride ion was placed in the most positive potential point on a surface enclosing the molecule at a distance of 0.7 nm from any heavy atom of the protein. The calculation and placement of ions was repeated until the global charge was zero.

Systems contained ca. 55,000 atoms.

The solute molecules, including ions, were fixed and the ensemble was energy minimized by 300 conjugate gradients steps using periodic boundary conditions and the particle mesh Ewald (PME) method for electrostatic interactions (6). PME employed a grid of 128 x 128 x 128 points. The PME tolerance was set to 10^{-6} that, together with the cutoff of 1.2 nm, resulted in an Ewald coefficient of 0.0257952 nm^{-1}. The system was then relaxed by 200 conjugate gradients minimization steps. The dielectric was set to 10.0 in order to minimize the effect of missing solvent and the cut-off was 1.2 nm. The minimized system was solvated using the module solvate in the VMD software package (7) in a box with margins at 0.5 nm distance from any solute atom.

Then, it was further relaxed, keeping the solute (including the ions) fixed, by molecular dynamics simulation. The system was heated to 300 K in 2 ps and a further 18 ps simulation was run in order to let

water molecules reorient, consistently with the average lifetime of a hydrogen bond in water (8). The system without restraints was energy minimized by 300 conjugate gradients minimization steps, then heated to 300 K in 2 ps and a further 118 ps simulation was run in order to let the system equilibrate after which the final production run could start.

FIGURES

Figure S1. ^{1}H-^{15}N resolved HSQC for the Pax-8 Prd Box domain recorded at 750 MHz.

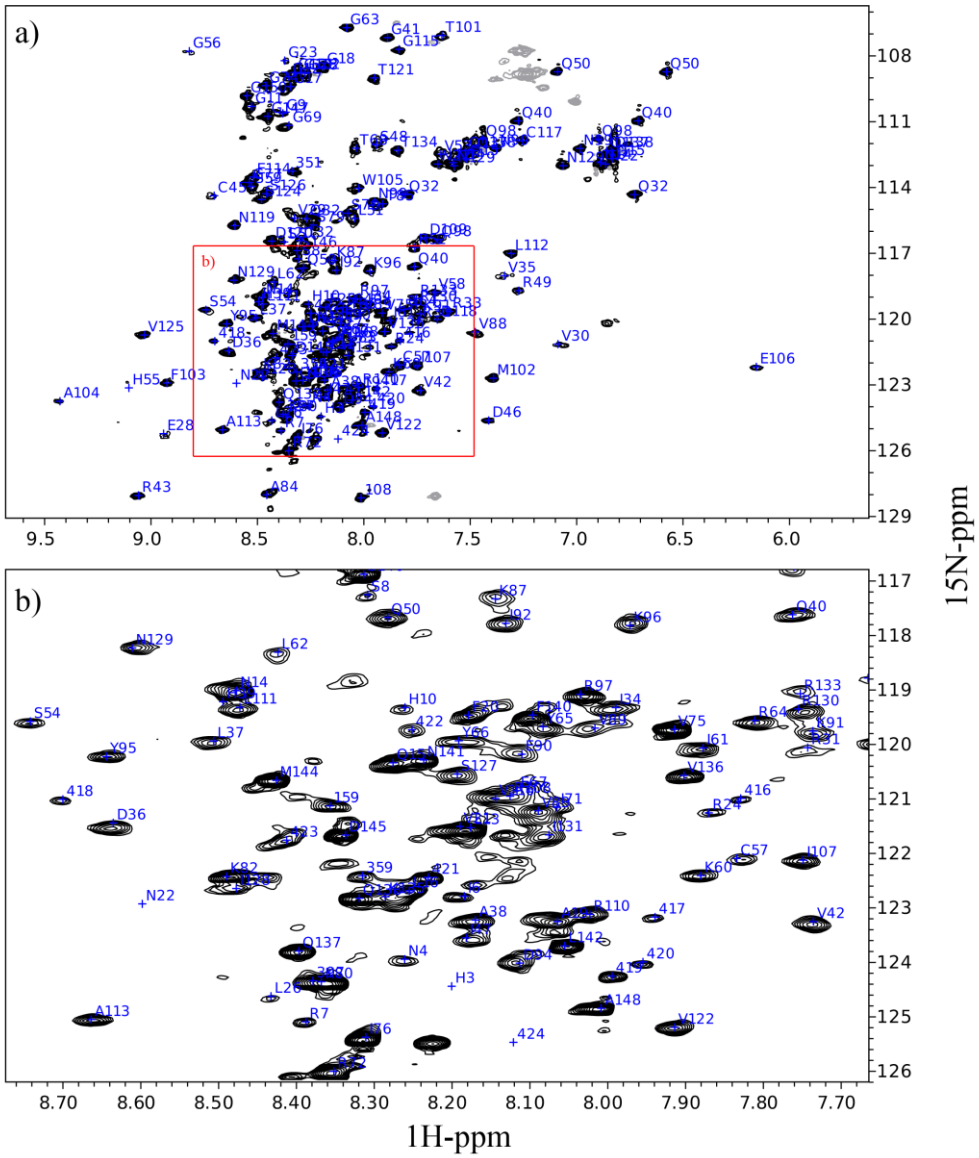


Figure S2. Experimental secondary-structure diagnostic connectivities. Boxes evidence final validated α-helices.

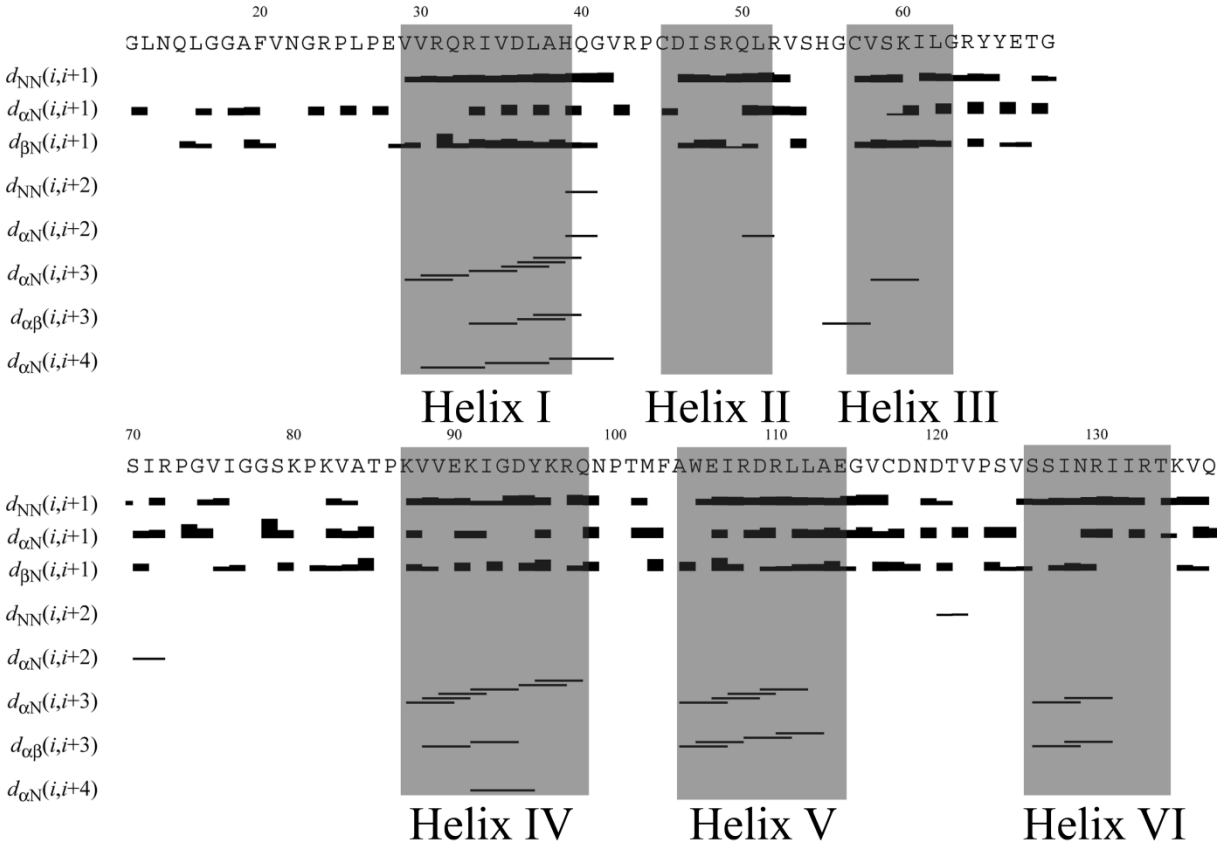


Figure S3. Control circular dichroism spectrum obtained at 298 K for the Pax-8 Prd Domain, a red line enlighten the θ value at 220 nm.

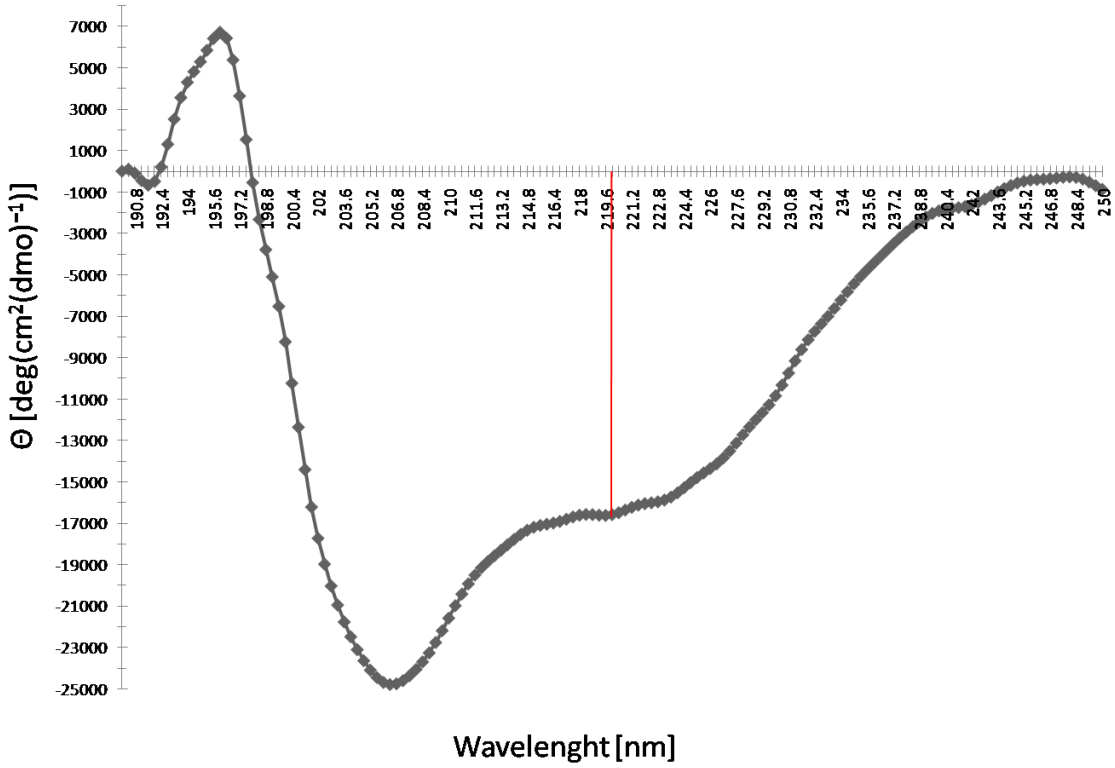


Figure S4. Long range restraints found in Pax-8 Prd Box Domain, red boxes evidence α-helical residues.

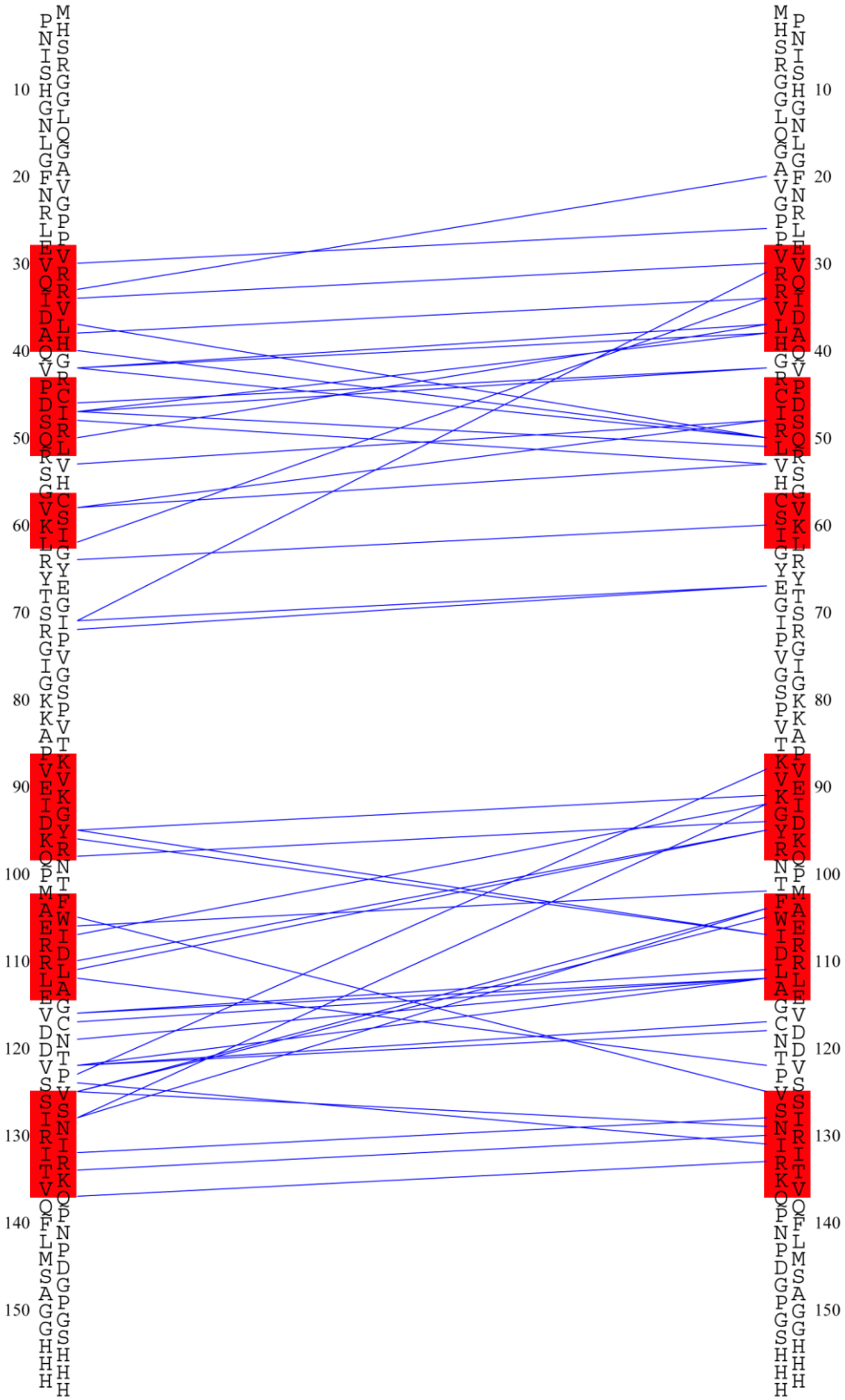


Figure S5. Random Coil Index a) and S2 values b) for the Pax-8 Paired Box Domain. Gray stripes indicate the α-helices found experimentally.

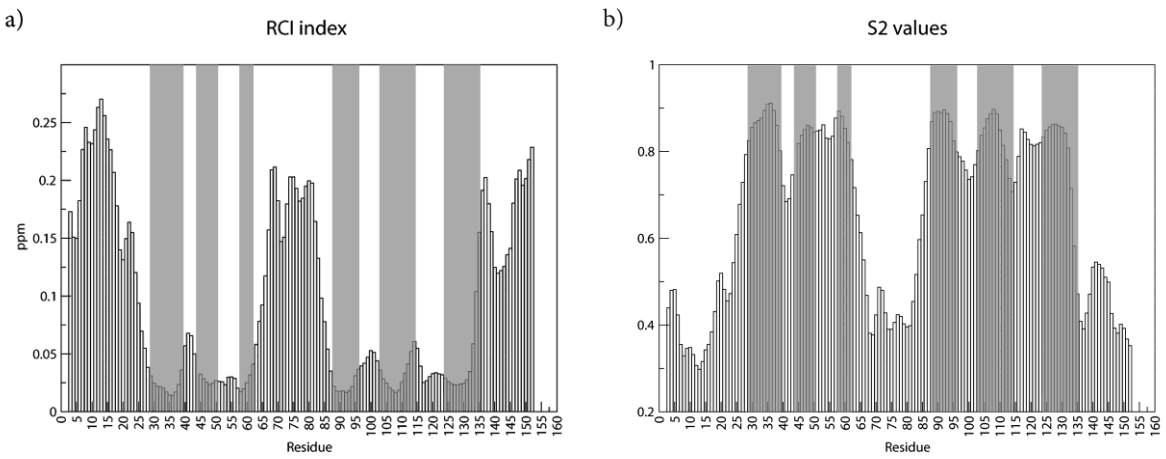
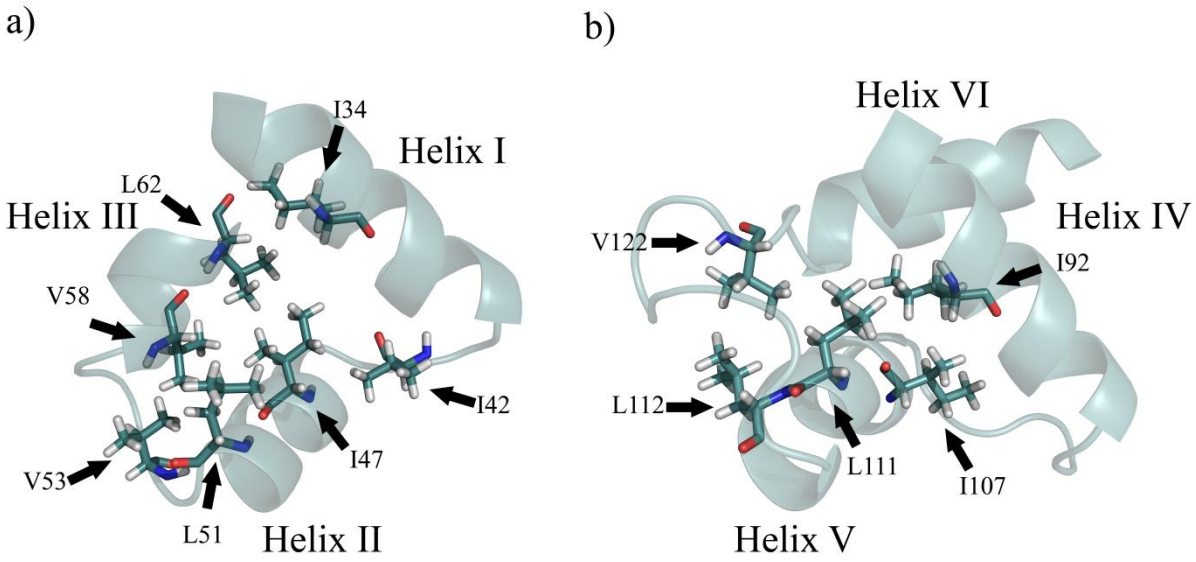


Figure S6. Important hydrophobic residues contributing to fold PAI and RED subdomains (in panels a and b, respectively).



TABLES

Table S1a. Hydrophobic interactions, supported by NOE data, found between the conserved residues in Pax Homologues

Table S1a.			**Conserved Hydrophobic Interactions**				
Residue		Residue		%	Pax-5	Pax-6	NOE
VAL	42	ILE	47	65	v	v	v
ILE	47	LEU	51	90	v	v	v
VAL	53	VAL	58	55	v	v	v
ILE	34	LEU	62	55		v	v
ILE	92	ILE	107	60	v	v	v
LEU	111	LEU	112	70	v		v
LEU	111	VAL	116	85	v	v	v
LEU	112	VAL	122	60	v	v	v

Table S1b. Salt bridges found in the Pax-8 ensemble, in evidence the conserved ones among the crystal homologues.

Table S1b.			**Conserved Salt Bridges**					
Residue			**Residue**			**%**	**Pax-5**	**Pax-6**
ASP	46	OD1	ARG	43	NH1	50	v	v
ASP	46	OD2	ARG	43	NH1	50	v	v
ASP	46	OD2	ARG	43	NH2	50	v	v
ASP	46	OD1	ARG	49	NH1	75	v	v
ASP	46	OD1	ARG	49	NH2	70	v	v
ASP	46	OD2	ARG	49	NH1	80	v	v
ASP	46	OD2	ARG	49	NH2	60	v	v
GLU	90	OE1	LYS	87	NZ	80		
GLU	90	OE2	LYS	87	NZ	70		
ASP	94	OD2	LYS	91	NZ	70		
ASP	94	OD1	ARG	97	NH1	85		
ASP	94	OD1	ARG	97	NH2	90		
ASP	94	OD2	ARG	97	NH1	65		
ASP	109	OD1	ARG	108	NH1	70	v	v
ASP	109	OD1	ARG	108	NH2	60	v	v
ASP	109	OD2	ARG	108	NH1	70	v	v
GLU	114	OE1	ARG	110	NH1	60	v	v
GLU	114	OE1	ARG	110	NH2	65	v	v
GLU	114	OE2	ARG	110	NH1	65	v	v
GLU	114	OE2	ARG	110	NH2	60	v	v

References

1. Bax, A., Clore, G. M. and Gronenborn, A. M. (1990). *J. Magn. Reson.* **88**, 425-431
2. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R. & Formankay, J. D. (1993) *J. Magn. Reson. Ser. B* **101**, 333-337
3. Berendsen, H. J. C., van der Spoel, D., & van Drunen R. (1995). *Comp. Phys. Comm.* **91**, 43-56.
4. Kale, L., Skeel, R., Bhandarkar , M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K. & Schulten K. (1999). *J. Comp. Phys.* **151**, 283-312.
5. Madura, J. D., Briggs, J. M., Wade, R., Davis, M. E., Luty, B. A., Ilin, A., Antosiewicz, J. A., Gilson, M. K., Bagheri, B., Ridgway Scott, L. & McCammon J.A. (1995). *Comput. Commun. Phys.* **91**, 57-95.
6. Darden, T., York, D. & Pedersen, L. (1993). *J. Chem. Phys.* **98**, 10089-10092.
7. Humphrey, W., Dalke, A., and Schulten, K. (1996) *J. Mol. Graph.* **14**, 33-38.
8. Keutsch, F., N., and Saykally, R. J. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10533-10540.