Supporting Information to:

**Mapping the Intramolecular Vibrational Energy Flow in Proteins Reveals Functionally Important Residues**

Leandro Martínez\(^1\,*\), Ana C. M. Figueira\(^2\), Paul Webb\(^3\), Igor Polikarpov\(^1\), Munir S. Skaf\(^4\,*\)

1. Institute of Physics of São Carlos, University of São Paulo. São Carlos, SP, Brazil.
2. National Laboratory of Biosciences, Brazilian Association for Synchrotron Light. Campinas, SP, Brazil.
3. Diabetes Center, The Methodist Hospital, Houston, TX, U.S.A.
4. Institute of Chemistry, State University of Campinas. Campinas, SP, Brazil.

**Materials and Methods**

*Molecular Dynamics Simulations*

Molecular dynamics simulations were performed with the crystal structures of TR\(\alpha\) and TR\(\beta\) bound to the natural ligand T3, determined by Polikarpov and co-workers (Nunes, 2004; Nascimento, 2006) and are refined to 1.87 and 2.4 Å, respectively. The structures were solvated with a 15Å thick water shell and one ion for each charged residue with Packmol (Martínez, 2003; Martínez, 2009a). Keeping all protein atoms fixed 1000 steps of Conjugate Gradient minimization were performed, followed by 100 ps MD with velocity rescaling at every picosecond to thermalize the solvent to 300K. Then, the constraints were removed from side-chain atoms and 100 ps MD with the same thermal coupling was performed. Finally, another 100 ps thermalization run with the same protocol was performed for the system without any restraint. From these equilibrated runs, 2 ns MD simulations were performed, from which 20 equally time-spaced frames were extracted for the thermal diffusion simulations. All minimization and equilibratons steps were performed with NAMD (Phillips, 2005) using CHARMM force field (MacKerell, 1998). Additional simulation details and ligand parameters can be found in our previous publications (Martínez, 2005, Martínez, 2006). A 2 fs time-step was used in all runs and
covalent bonds involving hydrogen atoms were kept rigid using RATTLE (Andersen, 1983).

Thermal diffusion simulations were performed following the general protocol described in (Ota & Agard, 2005) independently for each of the 20 frames extracted from the equilibrium TRα and TRβ simulations. From each frame, the solvent was removed and the structure was thermalized to 10K by performing a 5 ps MD with velocity rescaling at every step. For each 10K structure the following heating simulations were performed: 1) heating of each residue, independently, to 300K. Heating was performed by coupling an Langevin bath to all atoms of the heated residue with a coupling constant of 100 ps\(^{-1}\). Each heating simulations was 30 ps long. 2) The side chains of each residue in the structure were mutated to metil (thus, each residue was mutated to Alanine), and a similar simulation was performed. 3) The side chains of each residue were mutated to hydrogen (thus, each residue was mutated to Glycine) and the same simulation was performed. The LBDs of TRs contain about 260 residues. Therefore, 20 simulations were performed for native and 2 mutant structures for TRα and TRβ: a total of about 31,000 independent simulations were performed, or ~940 ns. Dispersion bars in final temperature plots are standard deviations of the 20 runs. In order to compute the temperatures of each residue in the structure velocity-trajectories were saved.

Additionally, the thermal response to ligand heating was studied in independent runs. In addition to the structures above, structures of TRα and TRβ LBDs bound to TRβ-selective ligands TRIAC (Martínez, 2009), GC-1 (Bleicher, 2008) and KB-141 (Ye, 2003) were used. The same equilibration protocols were used as above to obtain 20 initial LBD configurations for each ligand. The structures were thermalized to 10K and the response to ligand heating was evaluated by coupling the ligand atoms with a thermal bath at 300K with the same protocol as above. Therefore, additional 80 30 ps MD simulations comprise the set for ligand-heating.

*Side chain contributions and thermal diffusion maps*

The side-chain contributions for thermal energy transfer were evaluated by computing the difference of
the average final temperatures reached by the LBDs in native of mutant simulations, as represented in Figure 1 of the main manuscript. The maps of thermal diffusion represent the final temperature of each residue in response to the heating of every other residue. In the abscissa of each of these graphs we represent the “heated residue”, and in the ordinate the other residues of the protein. The color scales are linear in all graphics. For the thermal diffusion maps the color scale is from blue to red indicating larger temperatures.

*Time-dependence of the thermal response to ligand heating*

The time-dependence of thermal diffusion in the structure (main manuscript Figure 5c) was computed by recording the velocity-trajectory of all atoms in a typical 30 ps simulation of T3 heating in TRβ.

*Isotropic thermal response expectation to ligand heating*

The isotropic expectation of the temperatures were obtained as follows: For isotropic media, the expected temperature of any point in space is a function of the distance to the heat source according to a function of the form $T = A \exp\left[-B (d - d_o)\right] + C$ (Mikhailov, 1994). Therefore, given the distance of a residue to the heat source and knowing the values of the $A$, $B$, $d_o$ and $C$ constant, one could obtain the expected temperature for isotropic heat diffusion.

Therefore, for obtaining the isotropic heat diffusion equation that best fit to the obtained temperatures, we computed the shortest distance between atoms of the heated residue and the responding residue and fitted the isotropic diffusion equation, as a function of these distances, to the actual temperatures obtained. Since we are interested in the identification of the residues that respond anisotropically, it makes no sense fitting all data, since the anisotropic responses will be outliers of this fitting and should not be considered. The fitting was performed in order to obtain the best isotropic equation that adjusts the observed temperatures for 70% of the data. The outliers were automatically identified using a Low Order Value Optimization (LOVO) approach (Andreani, 2008) and the
optimization was performed using the GENCAN optimization algorithm (Birgin, 2002). 5000 different initial points (parameter values) were used for each fit in order to guarantee that the optimal fit and outlier identification were obtained.

Mutant TRβ and transactivation assays

Site-Directed Mutagenesis

The residues identified by ATD were submitted to site-directed mutagenesis. Primers designed to introduce single point mutations at the LBD of the nuclear receptor TRβ1 were used to amplify pCMV plasmids (Webb, 1995) encoding a wt construct of the human TRβ1 LBD. Single mutants were produced using QuikChange site-directed mutagenesis kit, according to the manufacture’s instructions (Stratagene). The presence of the mutation(s) was verified by DNA sequencing.

Cell Transactivation Assays and Transactivation

HeLa cells were seeded into 24-well plates at density of $1 \times 10^5$ cells/well and grown in 10% FBS-DMEM, under 95% air and 5% CO$_2$, at 37°C overnight, with 2 mM glutamine and 50 µg/ml streptomycin. The cells were then cotransfected with 10ng of pCMV-TRβ1 and with 100nM TRE (DR4) linked with Luciferase reporter. The plasmid pRL containing the Renilla luciferase gene was transfected simultaneously acting as the transfection control. TransFectinβ Lipid Reagent (Bio-Rad, Hercules, CA) was mixed with plasmids in DMEM and incubated at room temperature for 20 min. prior to adding to the culture media. The ratio of DNA (µg) to TransFectin (µl) was 1:3. The T3 ligand was subsequently added to the culture media 4 hours later and was incubated with the cells overnight. For activation assays, the ligand concentration in the cultures was $10^{-7}$M. The cell monolayer was then washed with Phosphate-buffered saline (PBS) and harvested with lysis buffer (Dual-Luciferase Report Assay system Promega, Madison, WI), following the manufacture’s instructions. Luciferase Activity
of the cell lysate was determined using the Luciferase Assay System (Promega, Madison, WI) and measured in a Safire2 luminescent counter (Tecan, Tecan US, NC, USA). The Renilla luciferase activity was measured using the same cell lysate (Dual-Luciferase Report Assay system Promega, Madison, WI) functioning as an internal control for the corresponding luciferase activity to adjust variation caused by transfection efficiencies. Luciferase assays were performed as described previously (Feng, 1998).
Figure S1. Sequence comparison between TRα and TRβ and their α-helical secondary structure. Red residues belong the binding pocket and dashed residues differ between subtypes. Residues are colored from N-terminal (red) to the C-terminal (blue). TRα Arginines 157, 176 and 188 are not conserved in TRβ. TRβ Arginine 391 is not conserved in TRα. The structural models used in the simulations start at TRα position 156.
Figure S2. TRβ LBD thermal response and side chain contributions for Glycine mutants.
Figure S3. TRα LBD thermal response and side chain contributions for Alanine mutants.
Figure S4. TRα LBD thermal response and side chain contributions for Glycine mutants.
Most pronounced side chain contributions

SCC ± SD / K

<table>
<thead>
<tr>
<th></th>
<th>TRβ (Ala mutants)</th>
<th>TRβ (Gly mutants)</th>
<th>TRα (Ala mutants)</th>
<th>TRα (Gly mutants)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Arg429 11.2 ± 2.2</td>
<td>Arg383 12.2 ± 2.9</td>
<td>Arg375 12.2 ± 2.1</td>
<td>Arg375 12.9 2.3</td>
</tr>
<tr>
<td>2</td>
<td>Arg383 10.9 ± 2.7</td>
<td>Arg429 12.1 ± 2.5</td>
<td>Arg189 10.0 ± 2.9</td>
<td>Arg189 9.9 2.9</td>
</tr>
<tr>
<td>3</td>
<td>Arg316 10.5 ± 2.4</td>
<td>Arg316 9.4 ± 2.3</td>
<td>Arg284 9.7 ± 2.3</td>
<td>Arg384 9.8 1.6</td>
</tr>
<tr>
<td>4</td>
<td>Arg410 8.3 ± 1.9</td>
<td>Arg282 8.5 ± 1.8</td>
<td>Arg356 9.4 ± 2.3</td>
<td>Arg356 9.6 2.5</td>
</tr>
<tr>
<td>5</td>
<td>Arg338 8.1 ± 1.7</td>
<td>Arg320 8.4 ± 1.9</td>
<td>Arg384 9.4 ± 2.0</td>
<td>Arg266 9.5 1.9</td>
</tr>
<tr>
<td>6</td>
<td>Arg243 8.0 ± 2.5</td>
<td>Arg410 8.0 ± 1.9</td>
<td>Arg266 9.1 ± 1.5</td>
<td>Arg284 9.2 2.3</td>
</tr>
<tr>
<td>7</td>
<td>Arg282 8.0 ± 2.0</td>
<td>Arg243 7.9 ± 2.1</td>
<td>Arg329 8.8 ± 2.6</td>
<td>Arg228 9.1 1.6</td>
</tr>
<tr>
<td>8</td>
<td>Arg320 7.9 ± 2.4</td>
<td>Lys342 7.5 ± 1.9</td>
<td>Arg262 8.7 ± 2.3</td>
<td>Arg329 8.9 2.1</td>
</tr>
<tr>
<td>9</td>
<td>Lys342 7.8 ± 2.0</td>
<td>Arg338 6.8 ± 2.7</td>
<td>Arg228 7.9 ± 1.7</td>
<td>Arg262 8.8 2.3</td>
</tr>
<tr>
<td>10</td>
<td>Arg438 6.9 ± 2.8</td>
<td>Phe354 6.8 ± 2.3</td>
<td>Lys304 7.0 ± 2.6</td>
<td>Gln320 8.2 1.7</td>
</tr>
<tr>
<td>11</td>
<td>Glu326 6.5 ± 2.5</td>
<td>Arg438 6.7 ± 2.8</td>
<td>Gln320 6.3 ± 1.9</td>
<td>Lys304 7.1 2.7</td>
</tr>
<tr>
<td>12</td>
<td>Asp367 6.4 ± 2.6</td>
<td>Gln374 6.5 ± 2.4</td>
<td>Asp272 6.1 ± 2.1</td>
<td>Tyr352 7.1 1.5</td>
</tr>
<tr>
<td>13</td>
<td>Phe354 6.1 ± 2.3</td>
<td>Glu369 6.6 ± 2.6</td>
<td>Tyr352 5.9 ± 2.1</td>
<td>Asp272 6.5 1.8</td>
</tr>
<tr>
<td>14</td>
<td>Lys288 5.7 ± 3.5</td>
<td>Glu326 6.3 ± 2.0</td>
<td>Lys366 5.9 ± 2.5</td>
<td>Arg157 6.2 1.8</td>
</tr>
<tr>
<td>15</td>
<td>Gln374 5.5 ± 2.2</td>
<td>Asp367 6.2 ± 3.1</td>
<td>Lys186 5.6 ± 2.3</td>
<td>Lys366 5.9 2.6</td>
</tr>
<tr>
<td>16</td>
<td>Tyr321 5.5 ± 3.1</td>
<td>Tyr321 5.8 ± 2.7</td>
<td>Arg157 5.4 ± 1.5</td>
<td>Tyr345 5.9 1.9</td>
</tr>
<tr>
<td></td>
<td>Residue</td>
<td>Value</td>
<td></td>
<td>Residue</td>
</tr>
<tr>
<td>---</td>
<td>---------</td>
<td>---------</td>
<td>---</td>
<td>---------</td>
</tr>
<tr>
<td>17</td>
<td>Arg391</td>
<td>5.3 ± 3.0</td>
<td></td>
<td>Tyr406</td>
</tr>
<tr>
<td>18</td>
<td>Tyr395</td>
<td>5.3 ± 2.0</td>
<td></td>
<td>Glu311</td>
</tr>
<tr>
<td>19</td>
<td>Leu377</td>
<td>5.3 ± 2.0</td>
<td></td>
<td>Leu315</td>
</tr>
<tr>
<td>20</td>
<td>Lys274</td>
<td>5.2 ± 1.6</td>
<td></td>
<td>Tyr395</td>
</tr>
</tbody>
</table>

**Table S1.** Most pronounced side-chain contributions to protein thermal response for TRβ and TRα, in which residues were individually mutated to Alanine or Glycine. Arginines that are conserved between subtypes are marked in blue.
Figure S5. Several Arginine residues interact with the protein inner structure; while other charged residues belong to the surface of the protein only, explaining why the non-bonded thermal diffusion is particularly important for Arginines. This suggests that charged residues belonging to the inner body of proteins are responsible for driving thermal dissipation from the protein cores. Stereo images are shown for the positions of (a) Arginines, (b) Glutamic Acids, (c) Lysines and (d) Aspartic acids.
Figure S6. Protein response to the heating of various TR ligands. T3 is the physiological hormone. TRIAC, GC1 and KB141 are TRβ selective ligands with pharmaceutical importance. Dispersion bars indicate the standard deviation of 20 runs performed on independent structure obtained from equilibrium MD simulations.
Figure S7. Thermal response profiles for (a) Peroxisome-Proliferated Activated Receptor-γ (pdb id. 1FM6) (Gampe, 2000) and (b) Hiperthermostable Xilanase (pdb id. 2VUJ) (Dumon, 2008). Arginines have also distinguished capacities to transfer heat to the structure in these models. Thermal responses for the hiperthermostable Xilanase are generally higher than those observed for NR LBDs.
References


hormone receptor isoforms.” BMC Struct. Biol. 8, 8, 2008.


(Martínez, 2006) L. Martínez, P. Webb, I. Polikarpov, M. S. Skaf, “Molecular Dynamics Simulations of Ligand Dissociation from Thyroid Hormone Receptors: Evidence of the Likeliest


(Taniyama 2001b) M. Taniyama, Y. Ban, N. Momotani, F. Makino, K. Ito, Y. Ban, “Three Japanese patients from two families with generalized resistance to thyroid hormone with mutations in


