

Fig. 1: amino acid composition of the N-terminal region (left panel) and remaining sequences (right panel) of type I NTR proteins. *Trypanosoma cruzi* values, presented as the average of all 11 sequences analysed, are black, and *T. rangeli* values are red (Tre), green (SC58), and blue (Choachi). *T. brucei*, presented as the average of two sequences, is pink, and the calcineurin B sequences from *T. cruzi* and *T. rangeli*, presented as averages, are grey.

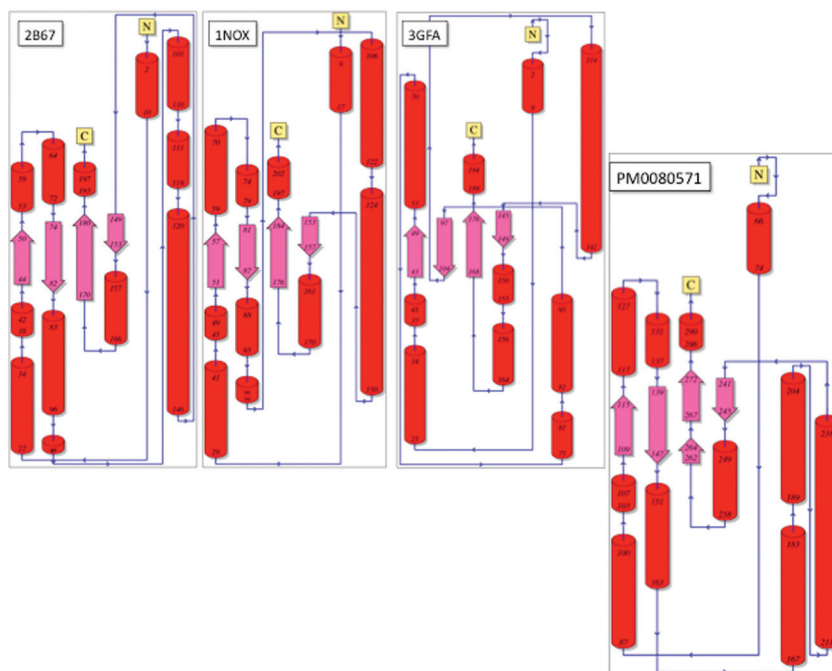
TABLE I  
Analysed type I NTR sequences with GenBank and Uniprot accession numbers

| Label       | Organism and classification         | GenBank      | Uniprot    |
|-------------|-------------------------------------|--------------|------------|
| Tc_058      | <i>Trypanosoma cruzi</i> (DTU I)    | AHI85558     | A0A075DCQ6 |
| Tr_Tre      | <i>Trypanosoma rangeli</i> (KP1(-)) | AHI85557     | A0A075DCB2 |
| Tr_SC58     | <i>T. rangeli</i> (KP1(-))          | ESL06666     | A0A061IXS5 |
| Tr_Choachi  | <i>T. rangeli</i> (KP1(+))          | AHI85556     | A0A075DD81 |
| Tc_CLB1     | <i>T. cruzi</i> (DTU VI)            | XP_810645    | Q4D8D9     |
| Tc_CLB2     | <i>T. cruzi</i> (DTU VI)            | XP_812226    | Q4DCW9     |
| Tc_FCH      | <i>T. cruzi</i> (DTU II)            | AFJ52880     | I6Q192     |
| Tc_YLY      | <i>T. cruzi</i> (DTU I/II)          | AFJ52881     | I6PZL1     |
| Tc_CG       | <i>T. cruzi</i> (DTU II)            | AFJ52882     | I6Q1A4     |
| Tc_MR       | <i>T. cruzi</i> (DTU II)            | AFJ52884     | I6PZV2     |
| Tc_W3534    | <i>T. cruzi</i> (DTU I)             | AFJ52889     | I6PZV5     |
| Tc_DA       | <i>T. cruzi</i> (DTU I)             | AFJ52890     | I6Q197     |
| Tc_Gal52    | <i>T. cruzi</i> (DTU I)             | AFJ52893     | I6Q086     |
| Tc_Sylvio   | <i>T. cruzi</i> (DTU I)             | EKG06918     | K4E8J0     |
| Tcm_B7      | <i>T. cruzi</i>                     | EKF39034     | K2NMK8     |
| Tv_Y486     | <i>Trypanosoma vivax</i>            | CCC49376     | G0TZH2     |
| Teg_IL3000  | <i>Trypanosoma congolense</i>       | CCD13725     | F9W935     |
| Tb_927      | <i>Trypanosoma brucei</i>           | XP_846343    | Q57XU3     |
| Tb_Dal972   | <i>T. brucei gambiense</i> DAL972   | XP_011775280 | C9ZUB5     |
| Lm_Friedlin | <i>Leishmania major</i>             | XP_001687543 | Q4QJD6     |
| Lmx_U1103   | <i>Leishmania mexicana</i>          | XP_003871956 | E9AKH2     |
| Li_JPCM5    | <i>Leishmania infantum</i> JPCM5    | XP_001463029 | A4HSI8     |
| Lb_M2904    | <i>Leishmania braziliensis</i>      | XP_001561879 | A4H4B0     |
| Ld_BPK282A1 | <i>Leishmania donovani</i>          | XP_003858253 | E9B8F6     |
| Lp_PSC-1    | <i>Leishmania panamensis</i>        | XP_010703828 | A0A088S1Z7 |
| G-lambli1   | <i>Giardia lamblia</i>              | XP_001705312 | A8BRN7     |
| G-lambli2   | <i>G. lamblia</i>                   | XP_001707931 | A8BCL1     |
| B. cereus   | <i>Bacillus cereus</i>              | AAP0997      | Q81BW8     |
| E. coli     | <i>Escherichia coli</i>             | NP_415110    | P38489     |
| Tc_CalB     | <i>T. cruzi</i> calcineurin B       | AEG15429     | S4TLQ2     |
| Tr_CalB     | <i>T. cruzi</i> calcineurin B       | AEG15431     | S4TLX5     |

The model was evaluated using PDBSum, and the experimental structure was compared with those of other nitroreductases (NTR) proteins.

| Plot statistics from Procheck based on Ramachandran plots |     |       |
|---|-----|-------|
| Residues in most favored regions                          | 202 | 81.1% |
| Residues in additional allowed regions                    | 39  | 15.7% |
| Residues in generously allowed regions                    | 4   | 1.6%  |
| Residues in disallowed regions                            | 4   | 1.6%  |

Topology diagram of secondary structure elements (SSE) from the PDBSum server



Helices are indicated as red cylinders, and beta sheets are indicated as pink arrows.

Summary of the PDBeFold results for comparison of the four structures

|                            |      |                        |      |
|----------------------------|------|------------------------|------|
| Number of aligned residues | 147  | Number of aligned SSEs | 7    |
| Overall RMSD               | 1.64 | Overall Q-score        | 0.36 |

Secondary structure alignment (SSE)

|                     |                  |
|---------------------|------------------|
| PM0080571 (Tre_NTR) | HHHShshhhhHs-HSh |
| 1NOX_A              | HHHShhs-hhHs-HSh |
| 3GFA_A              | HHHShh--hhHshHSH |
| 2B67_D              | HHHShhs-hhHs-HSh |

|      |         | Sequence identity (%) |        |        |        |
|------|---------|-----------------------|--------|--------|--------|
|      |         | Tre_NTR               | 1NOX_A | 3GFA_A | 2B67_D |
| RMSD | Tre_NTR |                       | 34.0   | 16.3   | 25.9   |
|      | 1NOX_A  | 1.59                  |        | 27.2   | 30.6   |
|      | 3GFA_A  | 1.86                  | 1.89   |        | 26.5   |
|      | 2B67_D  | 1.54                  | 1.39   | 1.46   |        |

Fig. 2: structural model validation.

Supplementary data

TABLE II  
Web servers used for subcellular localisation

| Server   | Web address   |
|--|---|
| WolfPsort                                      | <a href="http://www.genscript.com/wolf-psort.html">http://www.genscript.com/wolf-psort.html</a>                                 |
| PREDOTAR                                       | <a href="https://urgi.versailles.inra.fr/predotar/">https://urgi.versailles.inra.fr/predotar/</a>                               |
| MitoProt II                                    | <a href="https://ihg.gsf.de/ihg/mitoprot.html">https://ihg.gsf.de/ihg/mitoprot.html</a>   |
| PredSL   | <a href="http://aias.biol.uoa.gr/PredSL/">http://aias.biol.uoa.gr/PredSL/</a>   |
| PSORT II                                       | <a href="http://psort.hgc.jp">http://psort.hgc.jp</a>   |
| TargetP 1.1                                    | <a href="http://www.cbs.dtu.dk/services/TargetP/">http://www.cbs.dtu.dk/services/TargetP/</a>                                   |
| iPSort   | <a href="http://ipsort.hgc.jp">http://ipsort.hgc.jp</a>   |
| Eukmploc                                       | <a href="http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/">http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/</a>                 |
| ESLpred  | <a href="http://www.imtech.res.in/raghava/eslpred2/submit.html">http://www.imtech.res.in/raghava/eslpred2/submit.html</a>       |
| CELLO V.2.5 subCELLular Localization predictor | <a href="http://cello.life.nctu.edu.tw">http://cello.life.nctu.edu.tw</a>   |
| SubLoc v 1.0                                   | <a href="http://www.bioinfo.tsinghua.edu.cn/SubLoc/">http://www.bioinfo.tsinghua.edu.cn/SubLoc/</a>                             |
| YLoc   | <a href="http://abi.inf.uni-tuebingen.de/Services/YLoc/webloc.cgi">http://abi.inf.uni-tuebingen.de/Services/YLoc/webloc.cgi</a> |