



Available online at www.sciencedirect.com



COMMUNICATION

Structural Biology Sheds Light on the Puzzle of Genomic ORFans

Naomi Siew<sup>1,2,3</sup> and Daniel Fischer<sup>2,3\*</sup>

<sup>1</sup>Department of Chemistry  
Ben Gurion University  
Beer-Sheva 84105, Israel

<sup>2</sup>Bioinformatics Group,  
Department of Computer  
Science, Ben Gurion University  
Beer-Sheva 84105, Israel

<sup>3</sup>Buffalo Centre of Excellence in  
Bioinformatics/Computer  
Science and Engineering  
University of Buffalo, 901  
Washington St, Suite 300  
Buffalo, NY 14203, USA

Genomic ORFans are orphan open reading frames (ORFs) with no significant sequence similarity to other ORFs. ORFans comprise 20–30% of the ORFs of most completely sequenced genomes. Because nothing can be learnt about ORFans *via* sequence homology, the functions and evolutionary origins of ORFans remain a mystery. Furthermore, because relatively few ORFans have been experimentally characterized, it has been suggested that most ORFans are not likely to correspond to functional, expressed proteins, but rather to spurious ORFs, pseudo-genes or to rapidly evolving proteins with non-essential roles. As a snapshot view of current ORFan structural studies, we searched for ORFans among proteins whose three-dimensional structures have been recently determined. We find that functional and structural studies of ORFans are not as underemphasized as previously suggested. These recently determined structures correspond to ORFans from all Kingdoms of life, and include proteins that have previously been functionally characterized, as well as structural genomics targets of unknown function labeled as “hypothetical proteins”. This suggests that many of the ORFans in the databases are likely to correspond to expressed, functional (and even essential) proteins. Furthermore, the recently determined structures include examples of the various types of ORFans, suggesting that the functions and evolutionary origins of ORFans are diverse. Although this survey sheds some light on the ORFan mystery, further experimental studies are required to gain a better understanding of the role and origins of the tens of thousands of ORFans awaiting characterization.

© 2004 Elsevier Ltd. All rights reserved.

\*Corresponding author

Keywords: genomic ORFans; evolution; structural biology

Genomic ORFans<sup>1</sup> are orphan ORFs (open reading frames) that share no significant sequence similarity with any ORFs outside the genome where they reside (“singleton” and “paralogous” ORFans), or outside a set of closely related organisms (“orthologous” ORFans).<sup>2–4</sup> Genome sequencing of complete organisms has demonstrated that ORFans are integral components of most newly sequenced genomes;<sup>2,5,6</sup> their fraction in a newly sequenced genome is usually 20–30%,<sup>2,3</sup> and in

some cases, up to 60%.<sup>7</sup> Thus, ORFans are accumulating in the sequence databases at a rapid rate; over 30,000 singleton ORFans were counted in a recent census.<sup>8</sup> Because little can be inferred about the functions and structures of ORFans using standard bioinformatics tools, their abundance has been referred to as a “mystery”.<sup>3,9</sup> Furthermore, because relatively few ORFans have been studied experimentally,<sup>10–14</sup> many speculations regarding their roles and origins have been proposed. On the one hand, if ORFans correspond to expressed, functional proteins, they may be distant members of known protein families, with similar functions and three-dimensional (3D) structures, but with sequences that have diverged beyond recognition by standard sequence comparison tools.<sup>1,3</sup> In this case, knowing their 3D structures becomes essential in order to assign them to their corresponding families. Alternatively, ORFans may correspond to

Present address: Both authors, Buffalo Center of Excellence in Bioinformatics/Computer Science and Engineering, University at Buffalo, 901 Washington St. Suite 300, Buffalo, NY 14203, USA.

Abbreviations used: ORF, open reading frame; ORFans, orphan ORFs; FR, fold-recognition.

E-mail address of the corresponding author: dfischer@bioinformatics.buffalo.edu

64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126

novel proteins, unique to an organism or a lineage, with possibly new functions and/or 3D structures.<sup>1,3</sup> In either case, the mystery of why ORFans have no homologs remains. Are ORFans the result of rapid evolution,<sup>4,15,16</sup> of lateral gene transfer<sup>17</sup> from unknown organisms or are they the result of gene losses or of *de novo* generation?<sup>5,6,18–20</sup> On the other hand, it has also been suggested that most ORFans, especially the shorter ones, may correspond to non-essential, non-functional or non-expressed proteins.<sup>15,21–24</sup>

Here, as a snapshot view of current ORFan structural studies, we report a recent survey we conducted among newly determined 3D protein structures, and show how structural biology is already being essential in unraveling the ORFan puzzle.<sup>1,10,25</sup> We searched for ORFans among the PDB<sup>26</sup> entries released between June and December 2003 and found that out of the 172 protein chains sharing no significant sequence similarity to previously determined protein structures,<sup>27</sup> 17 correspond to ORFans and two correspond to “poorly conserved ORFs” or PCOs<sup>5,14</sup> (Table 1; for simplicity, in what follows we refer to these 19 proteins as

ORFans; see below). This strongly suggests that many ORFans correspond to real, foldable proteins, and not to sequencing errors or dead proteins. The relatively large percentage of ORFans among the newly determined structures (11%) suggests that ORFans may not be as underemphasized as previously suggested<sup>1,3</sup> and that experimental studies of ORFans have already become routine. In what follows, we show that these 19 ORFans provide interesting examples of the various types of ORFans.

Thirteen of these 19 newly determined ORFans correspond to proteins whose function was previously characterized experimentally (at least in the broad sense), and thus, are not “orphans” with regards to their functions. This suggests that many more ORFans with still unknown 3D-structures have already been characterized functionally. These 13 ORFans cover various functional categories, with at least five involved in transcription/translation, suggesting that, because of the high sequence divergence required in these processes, many ORFans may belong to these categories. The other six ORFans correspond to proteins of unknown

**Table 1.** The 19 ORFans with a recently determined 3D structure

PDB code	Organism	PDB description	Length (aa)	Has homologs in
1of5B <sup>28</sup>	<i>S. cerevisiae</i>	Mrna export factor Mex67-Mtr2	184	–
1oq1A	<i>B. subtilis</i>	Hypothetical protein Apc1120 (protein Yesu)	223	–
1mw5A	<i>H. influenzae</i>	Hypothetical protein Hi1480	187	–
1pp8U <sup>34</sup>	<i>T. vaginalis</i>	Initiator binding domain (Ibp39)	132	–
1q87A <sup>34</sup>	<i>T. vaginalis</i>	C-domain of the Inr binding protein	138	–
1q77A	<i>A. aeolicus</i>	Putative Universal Stress Protein (Hypothetical Protein Aq_178)	221	–
1rf8B <sup>38</sup>	<i>S. cerevisiae</i>	Translation initiation factor Eif4E	100	–
1nycA <sup>29</sup>	<i>S. aureus</i>	Staphostatin B	111	<i>S. warneri</i>
1oh1 <sup>39</sup>	<i>S. aureus</i>	Staphostatin A (hypothetical protein Sav1910)	109	<i>S. epidermidis</i>
1q8cA	<i>M. genitalium</i>	Conserved hypothetical protein (Mg027)	151	<i>M. pneumoniae</i>
1osyA <sup>40</sup>	<i>F. velupites</i>	Fip-Fve fungal immunomodulatory protein	115	<i>G. lucidum</i>
1r75A	<i>L. major</i>	Hypothetical protein	110	<i>T. brucei</i>
1nigA	<i>T. volcanium</i>	Hypothetical protein (Ta1238)	152	<i>T. acidophilum</i> , <i>F. acidarmanus</i>
1ofzA <sup>41</sup>	<i>A. auranta</i>	Fungal lectin	312	<i>A. oryzae</i> , <i>A. fumigatus</i>
1uf2P <sup>42</sup>	<i>Rice dwarf virus</i>	Rice dwarf virus (capsid protein)	421	Wound tumor virus, Rice gall dwarf virus
1q6aA <sup>43</sup>	<i>T. elongatus</i>	Circadian clock protein KaiA homolog	214	<i>T. erythraeum</i> , <i>Nostoc punctiforme</i> , <i>Nostoc</i> sp. PCC 9709, PCC 7120 & <i>P. membranacea</i> <i>Synechococcus</i> sp. PCC & WH, <i>Synechocystis</i> sp. PCC
1qv9A <sup>44</sup>	<i>M. kandleri</i>	Coenzyme F420-dependent Mtd	283	<i>M. jannaschii</i> , <i>M. mazei</i> , <i>M. acetivorans</i> , <i>M. barkeri</i> , <i>M. thermoautotrophicus</i> , <i>M. thermoautotrophicum</i> , <i>A. fulgidus</i>
1q5zA <sup>45</sup>	<i>S. typhimurium</i>	C-terminal actin binding domain of <i>Salmonella</i> invasion protein A (Sipa)	177	<i>S. enteritidis</i> , <i>S. enterica</i> Typhi, <i>C. violaceum</i> <sup>a</sup>
1n93A <sup>35</sup>	<i>Borna virus</i>	Nucleoprotein	375	<i>H. sapiens</i> , <i>M. musculus</i> <sup>b</sup>

ORFans are ORFs lacking significant sequence similarity with other ORFs, except for possibly ORFs from closely related organisms. We assess significant sequence similarity using the standard PSI-BLAST<sup>46</sup> sequence comparison tool until convergence, and the *e*-value threshold of 0.001. References for the PDB codes are included only if they were published before March 2004.

<sup>a</sup> Because of the presence of a homolog in *Chromobacterium violaceum*, 1q5zA is not a proper orthologous ORFan, but rather a poorly conserved ORF or PCO.

<sup>b</sup> 1n93A is also a PCO, which has probably been laterally transferred to humans and mice (see the text).

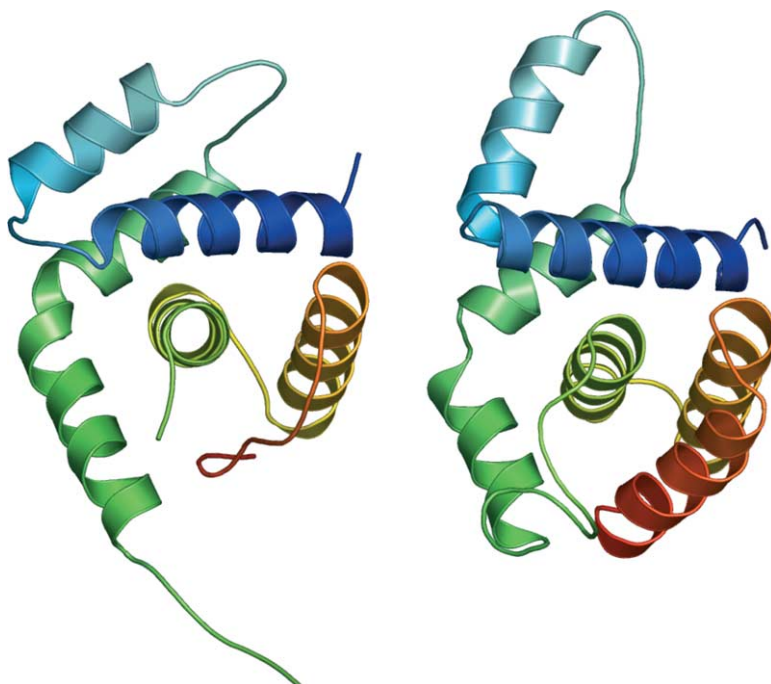
function (annotated as “hypothetical proteins”) whose structure was determined as part of structural genomics projects. Seven of the 19 ORFans correspond to singleton ORFans, and ten correspond to orthologous ORFans (have homologs only within closely related organisms). The 19 ORFans belong to organisms spanning all kingdoms: seven from Bacteria, three from Archea, seven from Eukarya and two viruses.

Despite the fact that ORFans show no significant sequence similarity to other proteins, the 3D structures of the majority of the ORFans clearly have previously observed folds. This suggests that they either correspond to highly divergent distant members of known protein families, with possibly similar functions or to proteins with unrelated functions whose structures have converged to a similar fold. For example, the essential messenger RNA export factor Mtr2 from yeast (1of5B), a singleton ORFan, was known to be similar in function to the metazoan p15 protein.<sup>28</sup> The previously determined 3D structure of the p15 protein revealed that it belongs to the NTF2-like family. The 3D structure of 1of5B revealed that Mtr2 is similar to that of p15, and thus, 1of5B represents a novel member of this family. This is a clear example of an ORFan with a highly divergent sequence, whose function and structure are similar to those of a known family. An example of an ORFan with a previously known fold, but with an unrelated function, is the bacterial virulence factor staphostatin B (1nycA), a cysteine protease inhibitor with a *Staphylococcus*-specific function.<sup>29</sup> Unexpectedly, its 3D structure turned out not to be similar to other

cystatins, but rather, to a variation of the lipocalin fold. This was a surprising result because staphylococci were not expected to contain lipocalin-like functions and no evidence of lipocalin-like properties has been identified in 1nycA. These examples illustrate that the 3D-structures of those ORFans whose broad function is known, can help to better understand their mechanisms of operation, and in some cases reveal their evolutionary origins.

The 3D-structures of the ORFans of unknown function, if similar to previously observed folds, can also be of help to generate verifiable hypotheses regarding their possible function and/or origin. One such example is the 3D-structure of the singleton ORFan of unknown function from *Aquifex aeolicus* (1q77A). Its structural similarity to the family of universal stress proteins lead to a putative functional assignment, which, if true, would imply that 1q77A is another highly divergent member of that family.

Could bioinformatics tools have predicted the approximate structures of those ORFans having previously observed folds? Fold-recognition (FR) methods,<sup>30,31</sup> applied without using the information from the recently released structures,<sup>27</sup> but using as templates previously determined structures, correctly predicted the structures of six of the 19 ORFans. Figure 1 shows the highly confident FR prediction of a *Mycoplasma*-specific hypothetical protein (1q8cA). This *M. genitalium* protein is an orthologous ORFan with a single homolog in *M. pneumoniae*. The FR result predicted that the 3D-structure of 1q8cA is similar to that of the transcription regulator NusB from *Mycobacterium*



**Figure 1.** A relatively accurate fold recognition prediction for *M. genitalium*'s conserved hypothetical protein MG027, 1q8cA. MG027 corresponds to an ORFan because with the exception of the close relative *M. pneumoniae*, it shows no sequence similarity to any other protein in the databases. Nevertheless, fold recognition is able to recognize the similarity between MG027 and a protein of known structure. The fold-recognition 3D model built without using the structural information from 1q8cA is shown on the right. The model, produced by the 3D-SHOTGUN fold-recognition method,<sup>36</sup> is based on the predicted structural similarity of MG027 with the previously released structure of the transcription regulator NusB from *Mycobacterium tuberculosis* (1evyA).<sup>32</sup> The sequence similarity between MG027 and 1evyA is only 13%. The

prediction is confirmed by the experimental structure of 1q8cA (left). Notice that no experimental data was observed for the loop region preceding the N-terminal two helices. The overall C-alpha RMSD between the predicted model and the experimental structures is 4.9 Å, with 105 residues superimposing with an RMSD of 2.5 Å.<sup>37</sup>

379 *tuberculosis* (1eyvA).<sup>32</sup> This prediction is now  
 380 confirmed by the experimental structure of 1q8cA.  
 381 Whether 1q8cA is also a transcription regulator or  
 382 not, remains to be verified experimentally.

383 Because no clear definition of what represents a  
 384 novel fold exists, it is difficult to tally the exact  
 385 number of ORFans with novel folds in our survey.  
 386 One accepted authority with regards to novel folds  
 387 is the SCOP<sup>33</sup> database. Unfortunately, the latest  
 388 SCOP release (1.65) only includes two of the 19  
 389 ORFans. Thus, we used structural comparison and  
 390 the available published reports to determine which  
 391 of the remaining ORFans correspond to novel folds.  
 392 We could identify only three cases of apparent  
 393 novel folds: 1q87A, 1q5zA and 1n93A. 1q87A  
 394 corresponds to the structure of the C-terminal  
 395 domain of the Inr binding protein from *Trichomonas*  
 396 *vaginalis*, which is of unknown function.<sup>34</sup> Another  
 397 interesting example of a novel fold is the nucleo-  
 398 protein from Born disease virus (1n93A).<sup>35</sup> This  
 399 virus is a member of the *Mononegavirales* that also  
 400 includes the measles and Ebola viruses. Because  
 401 homologs of this protein can also be found in  
 402 mammals (human and mouse), this is not a proper  
 403 ORFan, but a PCO,<sup>5,14</sup> probably illustrating an  
 404 example of lateral transfer, which could be one of  
 405 the mechanisms that has generated other ORFans.

406 The fact that the 3D structures of the majority of  
 407 ORFans in our survey correspond to previously  
 408 observed folds may be satisfying to some extent  
 409 because no new theories about their structural  
 410 origins are required. Many more 3D structures of  
 411 ORFans need to be solved in order to determine  
 412 whether ORFans will turn out to be an enriched  
 413 source of novel folds.

414 In summary, in our small survey we have found  
 415 examples of ORFans that are distant members of  
 416 known families, of ORFans with unknown function  
 417 or with lineage-specific functions and of ORFans  
 418 with novel, previously unseen folds. Although  
 419 many of the mysteries concerning ORFans remain  
 420 (i.e. their origins, their functions, their isolation in  
 421 sequence space), it is clear from this survey that  
 422 many ORFans correspond to real, functional, and in  
 423 some cases, essential, proteins. Surprisingly, 11% of  
 424 the newly determined structures that we con-  
 425 sidered corresponded to ORFans, suggesting that  
 426 ORFans may not be as underemphasized as  
 427 previously thought. Large-scale structure determi-  
 428 nations will be required to obtain further insights  
 429 about the evolution, the origin(s) and functions of  
 430 the tens of thousands of ORFans awaiting func-  
 431 tional and structural characterization.

432  
 433  
 434  
 435 **Additional material**

436  
 437 An expanded Table 1, containing various links to  
 438 the data is available†.

439  
 440 † [http://bioinformatics.buffalo.edu/ORFanage/](http://bioinformatics.buffalo.edu/ORFanage/3DORFans)  
 441 [3DORFans](http://bioinformatics.buffalo.edu/ORFanage/3DORFans)

**References**

1. Fischer, D. & Eisenberg, D. (1999). Finding families for genomic ORFans. *Bioinformatics*, **15**, 759–762.

2. Siew, N. & Fischer, D. (2003). Analysis of singleton ORFans in fully sequenced microbial genomes. *Proteins: Struct. Funct. Genet.* **53**, 241–251.

3. Siew, N. & Fischer, D. (2003). Twenty thousand ORFan microbial protein families for the biologist? *Structure (Camb)*, **11**, 7–9.

4. Malpertuy, A., Tekaiia, F., Casaregola, S., Aigle, M., Artiguenave, F., Blandin, G. *et al.* (2000). Genomic exploration of the hemiascomycetous yeasts: 19. Ascomycetes-specific genes. *FEBS Letters*, **487**, 113–121.

5. Siew, N. & Fischer, D. (2003). Unravelling the ORFan puzzle. *Comp. Funct. Genomics* 2003, 432–441.

6. Unger, R., Uliel, S. & Havlin, S. (2003). Scaling law in sizes of protein sequence families: from super-families to orphan genes. *Proteins: Struct. Funct. Genet.* **51**, 569–576.

7. Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W. *et al.* (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*, **419**, 498–511.

8. Siew, N., Azaria, Y. & Fischer, D. (2004). The ORFanage: an ORFan database. *Nucl. Acids Res.* **32**, D281–D283.

9. Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H. *et al.* (1996). Life with 6000 genes. *Science*, **546**, 563–567.

10. Fischer, D. (1999). Rational structural genomics: affirmative action for ORFans and the growth in our structural knowledge. *Protein Eng.* **12**, 1029–1030.

11. Monchois, V., Abergel, C., Sturgis, J., Jeudy, S. & Claverie, J. M. (2001). *Escherichia coli* ykfE ORFan gene encodes a potent inhibitor of C-type lysozyme. *J. Biol. Chem.* **276**, 18437–18441.

12. Alimi, J. P., Poirot, O., Lopez, F. & Claverie, J. M. (2000). Reverse transcriptase-polymerase chain reaction validation of 25 “orphan” genes from *Escherichia coli* K-12 MG1655. *Genome Res.* **10**, 959–966.

13. Brenner, S. E. (2001). A tour of structural genomics. *Nature Rev. Genet.* **2**, 801–809.

14. Shmueli, H., Dinitz, E., Dahan, I., Eichler, J., Fischer, D. & Shaanan, B. (2004). Poorly conserved ORFs in the genome of the archaea *Halobacterium* sp. NRC-1 correspond to expressed proteins. *Bioinformatics*, **20**, 1248–1253.

15. Schmid, K. J. & Aquadro, C. F. (2001). The evolutionary analysis of “orphans” from the *Drosophila* genome identifies rapidly diverging and incorrectly annotated genes. *Genetics*, **159**, 589–598.

16. Domazet-Loso, T. & Tautz, D. (2003). An evolutionary analysis of orphan genes in *Drosophila*. *Genome Res.* **13**, 2213–2219.

17. Daubin, V. & Ochman, H. (2004). Bacterial genomes as new gene homes: the genealogy of ORFans in *E. coli*. *Genome Res.* **14**, 1036–1042.

18. Wolfe, K. H. & Li, W. H. (2003). Molecular evolution meets the genomics revolution. *Nature Genet.* **33**, 255–265.

19. Ochman, H., Lawrence, J. G. & Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature*, **405**, 299–304.

20. Long, M. (2001). Evolution of novel genes. *Curr. Opin. Genet. Dev.* **11**, 673–680.

21. Skovgaard, M., Jensen, L. J., Brunak, S., Ussery, D. &

442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504

505 Krogh, A. (2001). On the total number of genes and  
 506 their length distribution in complete microbial genomes. *Trends Genet.* **17**, 425–428.  
 507  
 508 22. Amiri, H., Davids, W. & Andersson, S. G. (2003). Birth  
 509 and death of orphan genes in Rickettsia. *Mol. Biol. Evol.* **20**, 1575–1587.  
 510  
 511 23. Dujon, B. (1996). The yeast genome project: what did  
 512 we learn? *Trends Genet.* **12**, 263–270.  
 513  
 514 24. Mackiewicz, P., Kowalczyk, M., Gierlik, A., Dudek,  
 515 M. R. & Cebrat, S. (1999). Origin and properties of  
 516 non-coding ORFs in the yeast genome. *Nucl. Acids Res.* **27**, 3503–3509.  
 517  
 518 25. Watson, J. D., Todd, A. E., Bray, J., Laskowski, R. A.,  
 519 Edwards, A., Joachimiak, A. *et al.* (2003). Target  
 520 selection and determination of function in structural  
 521 genomics. *IUBMB Life*, **55**, 249–255.  
 522  
 523 26. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G.,  
 524 Bhat, T. N., Weissig, H. *et al.* (2000). The Protein Data  
 525 Bank. *Nucl. Acids Res.* **28**, 235–242.  
 526  
 527 27. Rychlewski, L., Fischer, D. & Elofsson, A. (2003).  
 528 LiveBench-6: large-scale automated evaluation of  
 529 protein structure prediction servers. *Proteins: Struct. Funct. Genet.* **53**, 542–547.  
 530  
 531 28. Fribourg, S. & Conti, E. (2003). Structural similarity in  
 532 the absence of sequence homology of the messenger  
 533 RNA export factors Mtr2 and p15. *EMBO Rep.* **4**, 699–  
 534 703.  
 535  
 536 29. Rzychon, M., Filipek, R., Sabat, A., Kosowska, K.,  
 537 Dubin, A., Potempa, J. & Bochtler, M. (2003).  
 538 Staphostatins resemble lipocalins, not cystatins in  
 539 fold. *Protein Sci.* **12**, 2252–2256.  
 540  
 541 30. Bujnicki, J. M., Elofsson, A., Fischer, D. & Rychlewski,  
 542 L. (2001). Structure prediction meta server. *Bioinformatics*, **17**, 750–751.  
 543  
 544 31. Ginalski, K., Elofsson, A., Fischer, D. & Rychlewski, L.  
 545 (2003). 3D-Jury: a simple approach to improve protein  
 546 structure predictions. *Bioinformatics*, **19**, 1015–1018.  
 547  
 548 32. Gopal, B., Haire, L. F., Cox, R. A., Jo Colston, M.,  
 549 Major, S., Brannigan, J. A. *et al.* (2000). The crystal  
 550 structure of NusB from *Mycobacterium tuberculosis*.  
 551 *Nature Struct. Biol.* **7**, 475–478.  
 552  
 553 33. Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia,  
 554 C. (1995). SCOP: a structural classification of proteins  
 555 database for the investigation of sequences and  
 556 structures. *J. Mol. Biol.* **247**, 536–540.  
 557  
 558 34. Schumacher, M. A., Lau, A. O. & Johnson, P. J. (2003).  
 559 Structural basis of core promoter recognition in a  
 560 primitive eukaryote. *Cell*, **115**, 413–424.  
 561  
 562 35. Rudolph, M. G., Kraus, L., Dickmanns, A., Eickmann,  
 563 M., Garten, W. & Ficner, R. (2003). Crystal structure of  
 564 the borna disease virus nucleoprotein. *Structure (Camb)*, **11**, 1219–1226.  
 565  
 566  
 567  
 568  
 569  
 570  
 571  
 572  
 573  
 574  
 575  
 576  
 577  
 578  
 579  
 580  
 581  
 582  
 583  
 584  
 585  
 586  
 587  
 588  
 589  
 590  
 591  
 592  
 593  
 594  
 595  
 596  
 597  
 598  
 599  
 600  
 601  
 602  
 603  
 604  
 605  
 606  
 607  
 608  
 609  
 610  
 611  
 612  
 613  
 614  
 615  
 616  
 617  
 618  
 619  
 620  
 621  
 622  
 623  
 624  
 625  
 626  
 627  
 628  
 629  
 630

Edited by M. Levitt

(Received 2 March 2004; received in revised form 9 June 2004; accepted 19 June 2004)