CMLS, Cell. Mol. Life Sci. 57 (2000) 851–858
1420-682X/00/010851-8 $ 1.50 + 0.20/0

CMLS Cellular and Molecular Life Sciences

Research Article

Three-dimensional model of human TIP30, a coactivator for HIV-1 Tat-activated transcription, and CC3, a protein associated with metastasis suppression

M. E. Baker[a,*], L. Yan[b] and M. R. Pear[b]

[a]Department of Medicine, 0823, University of California, San Diego, 9500 Gilman Drive, La Jolla (California 92093-0823, USA), Fax + 1 619 543 7069, e-mail: mbaker@ucsd.edu
[b]Molecular Simulations Inc., 9685 Scranton Road, San Diego (California 92121-3752, USA), e-mail: lly@msi.com,michaelpear@home.com

Received 26 January 2000; received after revision 2 March 2000; accepted 27 March 2000

Abstract. Human TIP30 is a cofactor that specifically enhances human immunodeficiency virus-1 (HIV-1) Tat-activated transcription. The sequence of TIP30 is identical to that of CC3, a protein associated with metastasis suppression. TIP30/CC3 is a member of the short-chain dehydrogenases/reductases (SDR) family. Of the several experimentally determined SDR structures, *Escherichia coli* uridine diphosphate (UDP) galactose-4 epimerase is most similar to TIP30/CC3. Because the direct sequence similarity between TIP30/CC3 and *E. coli* UDP galactose-4 epimerase is low, we used the transitive nature of homology and employed two *Aquifex aeolicus* proteins as intermediaries in the homology modeling process. Comparison of our structural model with that of known SDRs reveals that TIP30/CC3 contains several well-conserved features, including a $\beta\alpha\beta$ fold at the amino terminus, which we predict binds NADP(H). TIP30/CC3 contains characteristic motifs at the catalytic site of SDRs, including a serine, tyrosine, and lysine that are important in catalyzing hydride transfer between substrate and cofactor. We also predict that a unique 20-amino acid sequence found at the amino terminus is an α-helix. Because this region contains several positively and negatively charged amino acids, it may dock TIP30/CC3 to other proteins. Our structural model points to this α-helix and the SDR-like part of TIP30/CC3 for mutagenesis experiments to elucidate its role in HIV-1 Tat-activated transcription, metastasis suppression, and other cellular functions.

Key words. HIV-1; Tat; metastasis; TIP30; CC3.

Xiao et al. [1] identified human TIP30 as a cofactor that specifically enhances human immunodeficiency virus-1 (HIV-1) Tat-activated transcription. The sequence of human TIP30 is identical to that of human CC3, a protein associated with the suppression of metastasis in small-cell lung carcinomas [2]. Sequence analysis [3] shows that TIP30/CC3 is a member of the short-chain dehydrogenases/reductases (SDR) family [4]. This enzyme family has been extensively studied because it includes enzymes that regulate the concentrations of steroids, such as cortisol, estradiol and testosterone, as well as prostaglandins in humans [4–7]. Moreover, several three-dimensional (3D) structures of SDRs have been solved with X-ray crystallography [8–12]. Combined with mutagenesis studies [4, 13–16], these structures provide a good understanding of the catalytic mechanism of SDRs. A 3D crystal structure of TIP30/CC3 would be useful in designing mutagenesis studies

* Corresponding author.

A. UDP-galactose-4-epimerase (1XEL) aligned with *Aquifex aoelicus* 2983546.

```
1XEL         1 MRVLVTGGSGYIGSHTCVQLLQNGHDVIILDNLCNSKRSVLPVIERLGGK
2983546      1  KVLVTGGAGYIGSHVVKALGEKGYEVLIYDNLSTGN--------EWAVL

1XEL        51 HPTFVEGDIRNEALMTEILHDHAIDTVIHFAGLKAVGESVQKPLEYYDNN
2983546     42 YGKLVKADLADKETLRRVFEEFKPDAVMHFAAYIVVPESVKEPLKYYRNN

1XEL       101 VNGTLRLISAMRAANVKNFIFSSSATVYGDNPKIPYVESFPTGTPQSPYG
2983546     92 VVNTINLLEVMQEFGVNKFVFSSSAAVYGIPESIPVKEDAPLN-PINPYG

1XEL       151 KSKLMVEQILTDLQKAQPDWSIALLRYFNPVGAHPSGDMGEDPQGIPNNL
2983546    141 ETKATVERILRDLKNSGKDFNYVSLRYFNVAGADPEGKIG-FAYPNPTHL

1XEL       201 MPYIAQVAVGRRDSLAIFGNDYPTEDGTGVRDYIHVMDLADGHVVAMEKL
2983546    190 IIRAVKTAAGEFDRLEIYGTDYPTPDGTCIRDYIHVTDLAEAHILALEYL

1XEL       251 ANKPGVHIYNLGAGVGNSVLDVVNAFSKACGKPVNYHFAPRREGDLPAYW
2983546    240 FSGGKSEVLNCGYGHGYSVLEVVNAVKKVTGVDFKVVEAPRREGDPPALV

1XEL       301 ADASKADRELNWR-VTRTLDEMAQD
2983546    290 ADNKKIKRVLNWEPKYDDLEFIIKT
```

B. *Aquifex aoelicus* 2983546 aligned with *Aquifex aoelicus* 2982870.

```
2983546      1 KVLVTGGAGYIGSHVVKALGEKGYEVLIYDNLSTGNEWAVLYGKLVKA--
2982870      1 KVFITGATGFVGRHIVRELLNRGYEV--HAGVRNLSKLERLFGNQVKGYI

2983546     48 -DLADKETLRRVFEEFKPDAVMHFAAYIVVPESVKEPLKYYRNNVVNTIN
2982870     49 VNFDEKDSIREALGKVNPDFVIHLIG--ILYEEKKKGITFERVHYGHTKN

2983546     98 LLEVMQEFGVNKFVFSSSAAVYGIPESIPVKEDAPLNPINPYGETKATVE
2982870     97 LVEVSKGFNVKKFLFMSALGTH---------DEAP----SRYHQTKRWAE

2983546    148 RILRDLKNSGKDFNYVSLRYFNVAGADPEGKIGFAYPNPTHLIIRAVKTA
2982870    134 REVI---NSG--LNYTIFRPSIILG--PEQKLFFDM--------------

2983546    198 AGEFDRLEIYGTDYPTPDGTCIRDYIHVTDLAEAHILALEYLFSGGKSEV
2982870    163 ---YKITKYIPVVALPDFGNYQFQPVDVRDVACAYAEALKNPETDRK--I

2983546    248 LN-CGYGHGYSVLEVVNAVKKVTGVDFKVVEAPRREGDPPALVADNKKIK
2982870    208 YELCG-TKVVTFKELLADIFSYWDRKVLMIPMPKKLMYFAGLI-----VE
```

C. *Aquifex aoelicus* 2982870 aligned with human TIP30/CC3.

```
2982870:                     1  KVFITGATGFVGRHIVRELLNRG---YEVHA
TIP30/CC3    1 MAETEALSKLREDFRMQNKSVFILGASGETGRVLLKEILEQGLFSKVTLI

2982870:    29 GVRNLSKLERLFGNQVKGYIVNFDEKDSIREALGKVNPDFVIHLIGILYE
TIP30/CC3   51 GRRKLTFDEEAYKN-VNQEVVDFEKLDDYASAF--QGHDVGFCCLGTTRG

2982870:    79 EKKKGITFERVHYGHTKNLVEVSKGFNVKKFLFMSALGTHDEAPSRYHQT
TIP30/CC3   98 -KAGAEGFVRVDRDYVLKSAELAKAGGCKHFNLLSSKGADKSSNFLYLQV

2982870    129 KRWAEREVINSGLN-YTIFRPSIILGPEQKLFFDMYKITKYIPVVALPDF
TIP30/CC3  147 KGEVEAKVEELKFDRYSVFRPGVLLCDRQESRPGEWLVRKLFG--SLPDS

2982870:   178 GNYQFQPVDVRDVACAYAEALKNPETDRKIYELCGTKVVTFKELLADIFS
TIP30/CC3  195 -WARGHSVPVVTVVRAMLNNVVRP-RDKQMELLENKAIHDLGKAHGSLKP
```

Figure 1. Pairwise alignments used to construct the 3D model of TIP30/CC3. (*A*) *E. coli* UDP galactose-4 epimerase (PDB code 1XEL) aligned with *A. aoelicus* 2983546. The root mean square deviation for the resulting 3D model is 0.50 Å over the entire alignment. (*B*) *A. aoelicus* 2983546 aligned with *A. aeolicus* 2982870. The root mean square deviation for the resulting 3D model is 0.70 Å for the first 237 amino acids, and 1.2 Å for 266 amino acids. (*C*) *A. aeolicus* 2982870 aligned with TIP30/CC3. The root mean square deviation for the resulting 3D model is 0.82 Å for the first 201 amino acids, and 1.30 Å for the first 216 amino acids.

to elucidate how TIP30/CC3 binds to Tat and stimulates transcription, as well as the role of TIP30/CC3 in metastasis suppression and other cellular functions. However, crystallization of TIP30/CC3 is uncertain and even if successful, the entire process leading to an X-ray structure is time consuming. A more rapid approach is to construct a 3D model of TIP30/CC3 using as a template a structure that has been solved by X-ray crystallography. This approach is feasible for TIP30/CC3 because although it is distantly related to those SDRs with known 3D structures, even distantly related proteins conserve important structural features [17–19].

Here we describe a 3D model of TIP30/CC3, which provides information about its structure, including a putative binding site for NADP(H). In addition, we have modeled residues 1–20, which are not homologous to any known protein in GenBank, as an α-helix with positively and negatively charged residues that may be important in stabilizing TIP30/CC3 binding to Tat or other proteins. If this region of TIP30/CC3 is involved in Tat binding, its electrostatic properties would be targets for therapeutics to regulate Tat-activated transcription of HIV-1.

Materials and methods

TIP30/CC3 is 242 amino acids long, of which the first 20 amino acids are not related to SDRs. Among the SDRs with experimentally determined structures, TIP30/CC3 is most similar to *Escherichia coli* uridine diphosphate (UDP) galactose-4 epimerase (PDB code 1XEL) [3], which we selected as a template to model TIP30/CC3. UDP galactose-4 epimerase is 342 amino acids long. However, many SDRs, including hydroxysteroid dehydrogenases, are only 240–250 residues long, indicating that a shorter segment can suffice for catalysis. This 250-residue segment on SDRs aligns with the amino-terminal region of UDP galactose-4 epimerase [4, 20] and TIP30/CC3. The remaining carboxy-terminal region of UDP galactose-4 epimerase is involved in substrate binding that is specific to this enzyme. Thus, our goal was to model residues 21–242 of TIP30/CC3 on the core SDR catalytic structure, which contains the nucleotide cofactor-binding domain near the amino terminus and the catalytic site, usually 140–160 amino acids from the amino terminus. The carboxy-terminal region of SDRs is important in specific activities of each enzyme and shows the most divergent amino acid sequences [4–7].

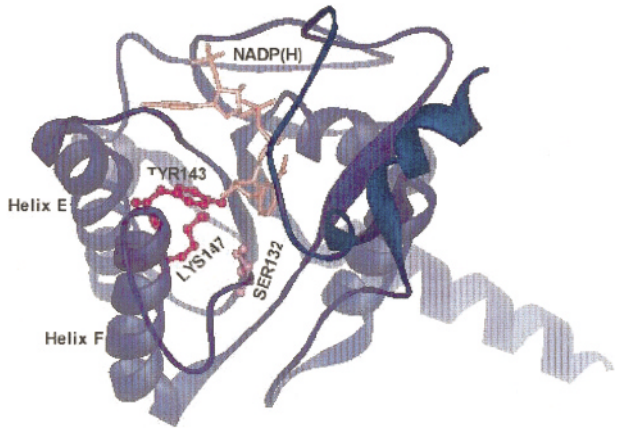

Figure 2. Structure model of TIP30/CC3. Serine-132, tyrosine-143 and lysine-147, which correspond to the catalytically active residues in SDRs and NADPH are shown. The α-helices E and F normally involved in a dimer interface are labelled.

There are islands of conservation that correspond to the cofactor-binding site and the catalytic site in the alignment of TIP30/CC3 and *E. coli* UDP galactose-4 epimerase. However, the two proteins have only 15% sequence identity over a segment of 160 amino acids, which is too low for accurate modeling directly from the template. Instead of directly modeling TIP30/CC3 on UDP galactose-4 epimerase, we took advantage of the transitive nature of homology and constructed models of two SDRs to bridge the gap in sequence similarity between TIP30/CC3 and UDP galactose-4 epimerase. One SDR, *Aquifex aoelicus* 2983546, is 41% identical to the *E. coli* UDP galactose-4 epimerase over a length of 330 amino acids. Moreover, the two sequences have 14% conservative replacements. At this level of similarity, we can model the carbon-α backbone and the side chains of *A. aoelicus* 2983546 with confidence [17–19]. *A. aoelicus* 2983546 is 29% identical to *A. aeolicus* 2982870, an SDR [21], and this protein in turn is 27% identical to TIP30/CC3 over a 175-amino acid segment, and 24% identical over a 206-amino acid segment. Thus, this three-step procedure yields a good description of TIP30/CC3 and will be especially good for conserved segments, which are likely to be important for key activities, such as nucleotide cofactor binding and catalyzing hydride transfer. The pairwise sequence alignments used for constructing each 3D model are shown in figure 1, along with the root mean square deviation for each 3D structure.

The model was constructed with the Molecular Simulations (MSI) Insight II software with the Modeler [22] and Homology [23] options.

To construct a model of TIP30/CC3 with NADPH, we extracted NADPH from mouse lung carbonyl reductase (PDB code 1CYD) [24]. NADPH was overlayed with NADH in TIP30/CC3, which was included from the 1XEL template during the modeling of TIP30/CC3. NADPH was aligned to NADH by a root mean square best fit of five common atoms.

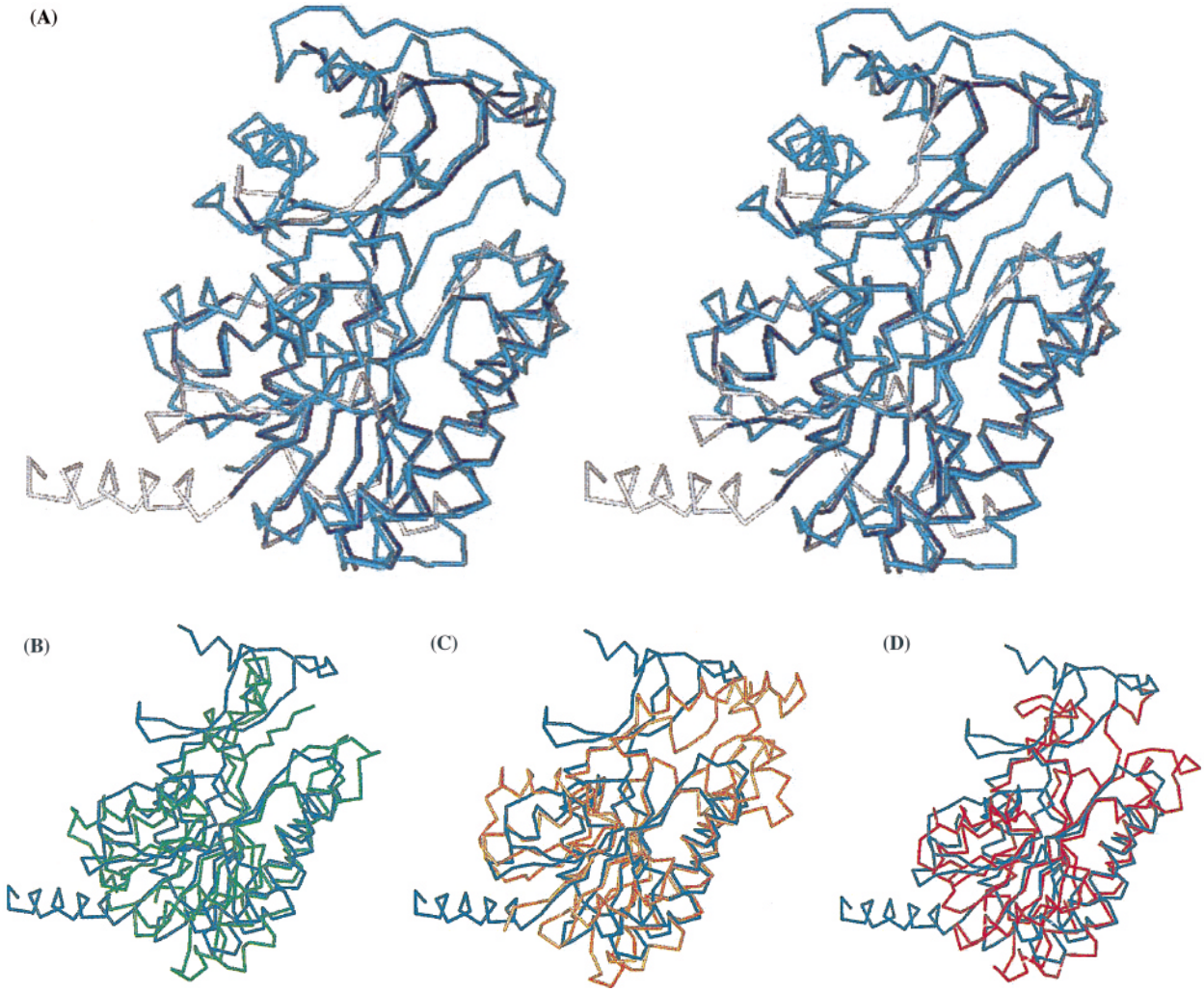

Figure 3. Overlay of TIP30/CC3 on 3D structures of several short-chain dehydrogenases/reductases. (*A*) The carbon-α chain of TIP30/CC3 is shown in stereo overlayed with *E. coli* UDP galactose-4 epimerase. The dark-blue chain shows 178 residues in TIP30/CC3 in which the root mean square deviation for the carbon-α chain with 1XEL is 0.82 Å. The less well conserved chain is in gray. 1XEL is in cyan. For the entire 220 residues of TIP30/CC3, the root mean square deviation is 1.59 Å. (*B–D*) TIP30/CC3 in blue overlayed with either human 17β-hydroxysteroid dehydrogenase, *E. coli* 7α-hydroxysteroid dehydrogenase, or *S. hydrogenans* 20β-hydroxysteroid dehydrogenase, respectively. There is good overlap for the α-helices and β-strands in TIP30/CC3 and these SDRs beginning at the amino terminus and continuing through α-helix F and β-strand F.

Results and discussion

The 3D model of TIP30/CC3 with NADPH shown in figure 2 contains α-helices A–F and the key catalytic sites for SDRs [4, 6–12]. An initial examination of residues 1–20 revealed two alanines, three glutamic acids, two leucines, and one methionine, all of which are strong α-helix formers by Chou-Fasman criteria [25]. Moreover, neither proline nor glycine, which are α-helix breakers, is present. To investigate further the secondary structure of residues 1–20, we used the GOR algorithm [26], which combines the Chou-Fasman analysis with other criteria to predict the secondary structure of peptides. GOR predicts that residues 1–20 form an α-helix, as shown in figure 2. This α-helix contains four amino acids with negatively charged side chains and three amino acids with positively charged side chains. As discussed later, these side chains could dock TIP30/CC3 to Tat or other proteins in the RNA polymerase II-containing complex.

To evaluate this model, we superimposed the carbon-α backbone of TIP30/CC3 on the template, *E. coli* UDP galactose-4 epimerase, as well as *E. coli* 7α-hydroxysteroid dehydrogenase, *Streptomyces hydrogenans* 20β-hydroxysteroid dehydrogenase, and human 17β-hydroxysteroid dehydrogenase as shown in figure 3a–d. The excellent superposition of the carbon-α backbone trace of TIP30/CC3 with known SDR struc-

tures indicates that we have a good model of the coenzyme-binding site and the catalytic site of TIP30/CC3. This is consistent with many studies showing that the 3D structure in homologues is conserved even when their amino acid sequences have less than 20% identity [17–19]. Interestingly, although the primary sequence of TIP30/CC3 is most similar to UDP galactose-4 epimerase [3], TIP30/CC3 lacks the extra loop of 12 amino acids between β-strand G and α-helix F that is found in *E. coli* UDP galactose-4 epimerase. In this respect, TIP30/CC3 is closer to hydroxysteroid dehydrogenases, which also lack this extra 12-amino acid segment [4–7].

Coenzyme specificity. The amino terminus of TIP30/CC3 contains the $\beta\alpha\beta$ motif that binds the nucleotide part of NAD(P)(H) of SDRs [4, 6, 8–12, 18] and other dehydrogenases [27, 28]. A short loop of three or four amino acids after the end of β-strand B often contains one or more amino acids that are spatially close to the adenine ribose and can confer specificity for NAD(H) or NADP(H) by either promoting or disfavoring binding. For example, a negatively charged amino acid such as aspartic acid will repel the ribose 2′-phosphate on NADP(H), favoring binding of NAD(H) [29–31].

We modeled NADPH in TIP30/CC3 and found a stabilizing coulombic attraction between the ribose 2′-phosphate and arginine-53 as shown in figure 4. This stabilizing interaction is similar to that found in many dehydrogenases that bind NADPH [31]. Moreover, lysine-54 also interacts with the ribose 2′-phosphate. TIP30/CC3 is unique in that none of the other SDRs analyzed by Mazza et al. [31] have an adjacent arginine/lysine pair that interacts with the 2′-phosphate. However, in 17β-hydroxysteroid dehydrogenase-type 1, arginine-37 and lysine-195 have a stabilizing coulombic attraction with the ribose 2′-phosphate in NADPH [31] indicating convergent evolution for this stabilizing interaction. We predict that NADP(H) is the coenzyme for TIP30/CC3.

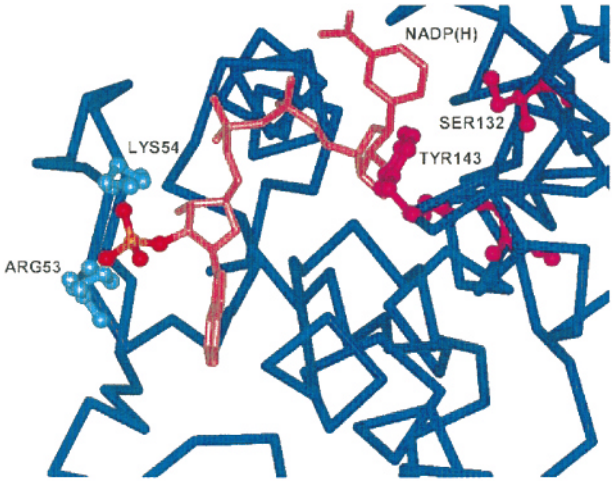

Figure 4. Interaction between NADPH and TIP30/CC3. Arginine-53 and lysine-54 have a stabilizing coulombic interaction with the ribose 2′-phosphate in NADPH.

Catalytic site. The catalytic site of SDRs contains a tyrosine that catalyzes hydride transfer and a lysine that forms hydrogen bonds with the 2′- and 3′-hydroxyl groups of the nicotinamide ribose. The tyrosine and lysine are separated by three residues (Tyr-Xaa-Xaa-Xaa-Lys) on α-helix F and are characteristic of SDRs [4–16]. A catalytically important serine is positioned upstream from the tyrosine at the end of β-strand G [4, 8–13]. TIP30/CC3 contains these three residues [3], which are shown in relationship to NADPH in figure 2. The SDR part of TIP30/CC3 contains only 220 residues, which is 20–30 residues shorter than mouse carbonyl reductase, *E. coli* 7α-hydroxysteroid dehydrogenase, and many other SDRs, excluding the epimerases that are over 300 residues long. The extra carboxy-terminal residues on epimerases bind either the UDP or GDP that is conjugated to sugar that undergoes the epimerase reaction.

Although TIP30/CC3 has the general structure and key amino acids found in the catalytic site of SDRs, we do not know if TIP30/CC3 is an enzyme. Nevertheless, a role for a catalytically active TIP30/CC3 in either Tat-activated transcription or metastasis suppression can be investigated by mutation of tyrosine-143 to phenylalanine, which would inactivate any SDR activity [4–7].

Dimer interface. Most SDRs are either dimers or tetramers [4, 16, 32, 34] that are stabilized by a four-α-helix bundle comprised of α-helices E and F from each subunit. The two F α-helices are closer to each other than are the two E α-helices. The core of this four-α-helix bundle within the SDR family is stabilized primarily by hydrophobic forces especially across α-helix F [8–12, 16, 32–34].

We modeled the dimer interface between the F α-helices in TIP30/CC3 using UDP galactose-4 epimerase and refined the interface by further alignment with that of *E. coli* 7α-hydroxysteroid dehydrogenase, which is a better template for this region of TIP30/CC3. As shown in figure 5, α-helix F contains several charged residues (Glu-149, Lys-153, and Glu-156) on the surface that would form the dimer interface. The coulombic interactions of these charged residues do not stabilize the dimer interface. Instead, Glu-149 in each subunit is separated from its counterpart by only 3.0 Å and, therefore, is a source of coulombic repulsion. Glu-149 is at an anchor point in the dimer interface that is sensitive to side chain interactions [16, 32, 34]. A mutation at the corresponding residue position in *Drosophila* alco-

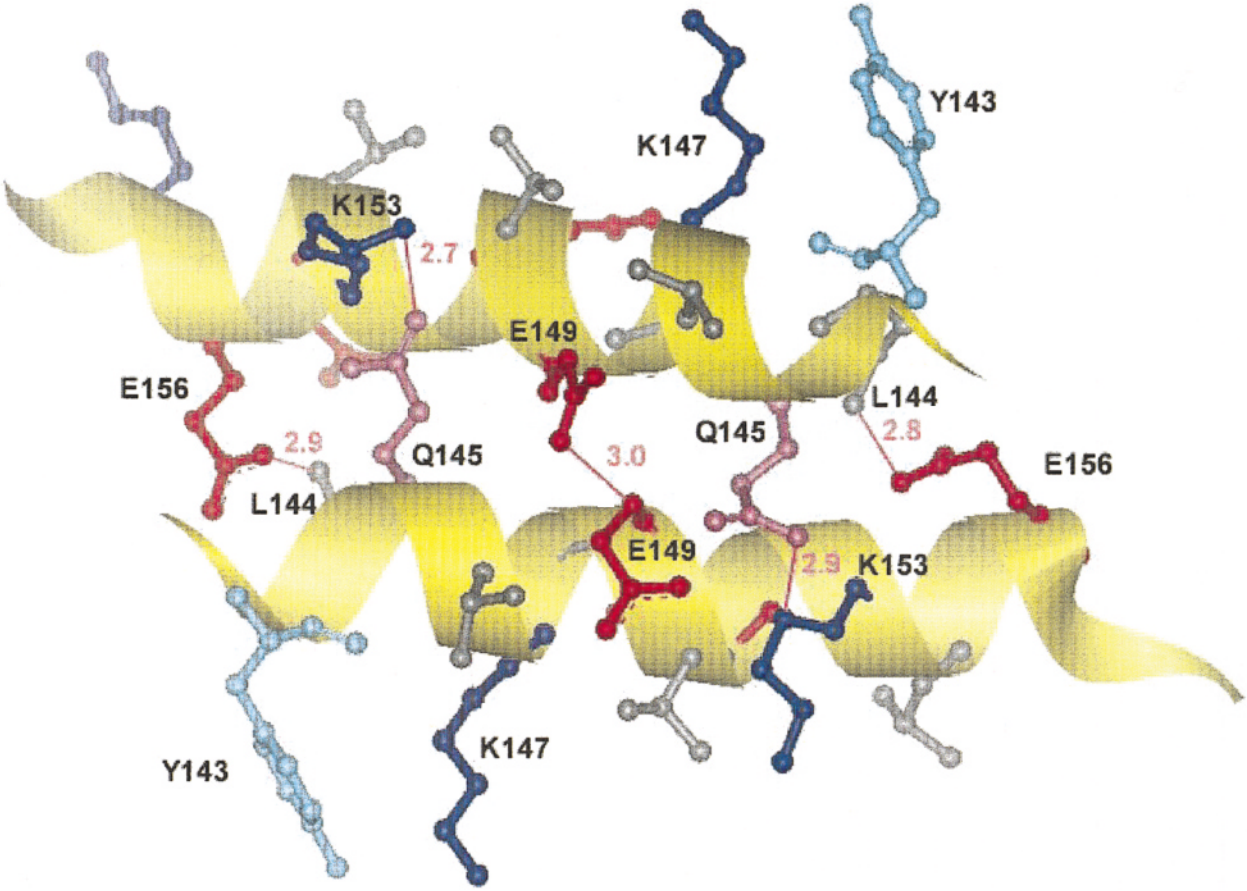

Figure 5. Dimer interface of α-helix F in TIP30/CC3. Glutamic acid-149 on each helix is about 3 Å from the other and they exert coulombic repulsion. Glutamine-145 and leucine-144, which are between the catalytic tyrosine-143 and lysine-147, are close to lysine-153 and glutamic acid-156, respectively. There are no stabilizing coulombic or hydrophobic interactions across α-helix F (E, glutamic acid; K, lysine; L, leucine; Q, glutamine; Y, tyrosine).

hol dehydrogenase destabilizes the dimer interface and leads to a loss of catalytic activity [34].

Most SDRs have amino acids with a small side chain in the segment between the catalytic tyrosine and lysine, so as not to interfere with the conformation of the catalytic residues. TIP30/CC3 also is unusual in having amino acids with long side chains between the catalytic Tyr-143 and Lys-147. TIP30/CC3 has Gln-145 only 2.9 Å from lysine-153 and Leu-144 only 2.8 Å from Glu-156, which could constrain movement of Tyr-143 and Lys-147.

These constraining interactions and the absence of stabilizing forces from either hydrophobic or coulombic interactions leads us to propose that TIP30/CC3 does not form a dimer like other SDRs.

With the exception of human and pig carbonyl reductase [35, 36], all known SDRs are dimers. The formation of a dimer is a means of stabilizing the catalytic tyrosine and lysine on α-helix F. In the two exceptional cases where no dimer is formed, an extra internal segment has been proposed to bind to α-helix F [16, 32] and stabilize the catalytic tyrosine and lysine on the inner surface by forming the equivalent of the dimer interface.

If TIP30/CC3 does form a dimer, the interface region must be very different from the typical SDR dimer interface. The charged properties of α-helix F on TIP30/CC3 suggest that it is involved in a protein-protein interaction, which could stabilize the configuration of tyrosine-143 and lysine-147. It is also possible that the first 20 residues at the amino terminal serve as a stabilizing helix. However, the position of this region does not appear to be sufficiently near α-helix F for an intramolecular interaction.

Uniqueness of residues 1–20 of TIP30/CC3. Shtivelman [2] noted that CC3 had strong sequence similarity to

open reading frames in *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, and *E. coli*. Despite sequence identity of over 41% between human TIP30/CC3 and its *C. elegans* homologue, the latter protein lacks the amino terminal 19 residues, as does the *S. cerevisiae* homologue.

Although, the *E. coli* homologue contains a 20-residue segment at its amino terminus, this segment has no similarity to that in TIP30/CC3. Moreover, the *E. coli* segment contains three glycines, which are not usually found in α-helices. Thus, our proposed helix on human and mouse TIP30/CC3 is unique to mammals and is likely to have some important cellular function.

Biological interactions of TIP30/CC3. Although TIP30/CC3 is known to bind to residues 1–48 of Tat [1], the details of the interaction between these two proteins have not been determined. Residues 1–48 on Tat contain six positively charged and three negatively charged amino acids, which could have coulombic interactions with either residues 1–20 or with α-helix F of TIP30/CC3 or both. Mutagenesis studies to remove or modify residues 1–20 on TIP30/CC3 would determine whether this segment binds Tat and whether it is important in mediating increased transcription of HIV by Tat.

At this time, the most clearly defined property of TIP30/CC3 is its binding of residues 1–48 of Tat. However, Shtivelman [2] has shown that TIP30/CC3 is important in suppression of metastasis of small-cell lung carcinoma. Thus, the interaction of TIP30/CC3 with proteins other than Tat is of interest for understanding the cellular function of TIP30/CC3 and its role in suppressing metastasis. When such proteins are found, mutations of TIP30/CC3 in tyrosine-143 and the amino-terminal α-helix should be helpful.

1 Xiao H., Tao Y., Greenblatt J. and Roeder R. G. (1998) A cofactor, TIP30, specifically enhances HIV-1 Tat-activated transcription. Proc. Natl. Acad. Sci. USA **95:** 2146–2151
2 Shtivelman E. (1997) A link between metastasis and resistance to apoptosis of variant small cell lung carcinoma. Oncogene **14:** 2167–2173
3 Baker M. E. (1999) TIP30, a cofactor for HIV-1 Tat-activated transcription, is homologous to short-chain dehydrogenases/reductases. Curr. Biol. **9:** R471
4 Jornvall H., Persson B., Krook M., Atrian S., Gonzalez-Duarte R., Jeffrey J. et al. (1995) Short-chain dehydrogenases/reductases (SDR). Biochemistry **34:** 6003–6013
5 Krozowski Z. (1992) 11β-Hydroxysteroid dehydrogenase and the short chain alcohol dehydrogenase (SCAD) superfamily. Mol. Cell. Endocrinol. **84:** C25–C31
6 Bailey T. L., Baker M. E. and Elkan C. P. (1997) An artificial intelligence approach to motif discovery in protein sequences: application to steroid dehydrogenases. J. Steroid Biochem. Mol. Biol. **62:** 29–43
7 Persson B., Krook M. and Jornvall H. (1991) Characteristics of short-chain alcohol dehydrogenase related enzymes. Eur. J. Biochem. **200:** 537–543
8 Varughese K. I., Xuong N. H., Kiefer P. M., Matthews D. A. and Whiteley J. M. (1994) Structural and mechanistic characteristics of dihydropteridine reductase: a member of the Tyr-$(Xaa)_3$-Lys-containing family of reductases and dehydrogenases. Proc. Natl. Acad. Sci. USA **91:** 5582–5586
9 Ghosh D., Wawrzak Z., Weeks C. M., Duax W. L. and Erman M. (1994) The refined three-dimensional structure of 3α,20β-hydroxysteroid dehydrogenase and possible roles of the residues conserved in chart-chain dehydrogenases. Structure **2:** 629–640
10 Tanaka N., Nonaka T., Tanabe T., Yoshimoto T., Tsuru D. and Mitsui Y. (1996) Crystal structures of the binary and ternary complexes of 7α-hydroxysteroid dehydrogenase from *Escherichia coli*. Biochemistry **35:** 7715–7730
11 Thoden J. B., Frey P. A. and Holden H. M. (1996) Crystal structures of the oxidized and reduced forms of UDP-galactose 4-epimerase isolated from *Escherichia coli*. Biochemistry **35:** 2557–2566
12 Breton R., Housset D., Mazza C. and Fontecilla-Camps J. C. (1996) The structure of a complex of human 17β-hydroxysteroid dehydrogenase with estradiol and $NADP^+$ identifies two principal targets for the design of inhibitors. Structure **4:** 905–915
13 Baker M. E. (1994) Sequence analysis of steroid- and prostaglandin-metabolizing enzymes: application to understanding catalysis. Steroids **59:** 248–258
14 Obeid J. and White P. C. (1992) Tyr-179 and Lys-183 are essential for enzymatic activity of 11β-hydroxysteroid dehydrogenase. Biochem. Biophys. Res. Commun. **188:** 222–227
15 Chen Z., Jiang J. C., Lin Z. G., Lee W. R., Baker M. E. and Chang S. H. (1993) Site-specific mutagenesis of *Drosophila* alcohol dehydrogenase: evidence for involvement of tyrosine-152 and lysine-156 in catalysis. Biochemistry **32:** 3342–3346
16 Tsigelny I. and Baker M. E. (1995) Structures important in mammalian 11β and 17β-hydroxysteroid dehydrogenases. J. Ster. Biochem. Mol. Biol. **55:** 589–600
17 Chothia C. and Lesk A. M. (1986) The relation between the divergence of sequence and structure in proteins. EMBO J. **5:** 823–826
18 Branden, C. and Tooze, J. (1991) Introduction to protein structure, Garland, New York
19 Chung S. Y. and Subbiah S. (1996) A structural explanation for the twilight zone of protein sequence homology. Structure **4:** 1123–1127
20 Baker M. E. and Blasco R. (1992) Expansion of the mammalian 3β-hydroxysteroid dehydrogenase/plant dihydroflavonol reductase superfamily to include a bacterial cholesterol dehhydrogenase, a bacterial UDP-galactose-4-epimerase, and open reading frames in vaccinia virus and fish lymphocystis disease virus. FEBS Lett. **301:** 89–93
21 Baker M. E., Grundy W. N. and Elkan C. P. (1999) A common ancestor for a subunit in the mitochondrial proton-translocating NADH:ubiquinone oxidoreductase (complex I) and short-chain dehydrogenases/reductases. Cell. Mol. Life Sci. **55:** 450–455
22 Sali A. and Blundell T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. **234:** 779–815
23 Homology (1998) User manual, Molecular Simulations, San Diego
24 Tanaka N., Nonaka T., Nakanishi M., Deyashiki Y., Hara A. and Mitsui Y. (1996) Crystal structure of the ternary complex of mouse lung carbonyl reductase at 1.8 Å resolution: the structural origin of coenzyme specificity in the short-chain dehydrogenase/reductase family. Structure **4:** 33–45
25 Chou P. Y. and Fasman G. D. (1978) Empirical predications of protein structure. Annu. Rev. Biochem. **47:** 251–276
26 Gibrat J. F., Garnier J. and Robson B. (1987) Further developments of protein secondary structure prediction using information theory: new parameters and consideration of residue pairs. J. Mol. Biol. **198:** 425–443
27 Wierenga R. K., De Maeyer M. C. and Hol W. G. J. (1985) Interaction of pyrophosphate moieties with α-helixes in dinucleotide binding proteins. Biochemistry **24:** 1346–1357
28 Wierenga R. K., Terpstra P. P. and Hol W. G. J. (1986) Prediction of the occurrence of the ADP-binding $\beta\alpha\beta$-fold in

proteins using an amino acid sequence fingerprint. J. Mol. Biol. **187:** 101–107
29 Chen Z., Tsigelny I., Lee W. R., Baker M. E. and Chang S. H. (1994) Adding a positive charge at residue 46 of *Drosophila* alcohol dehydrogenase increases cofactor specificity for $NADP^+$. FEBS Lett. **356:** 81–85
30 Nakanishi M., Matsuura K., Kaibe H., Tanaka N., Nonaka T., Mitsui Y. et al. (1997) Switch of coenzyme specificity of mouse lung carbonyl reductase by substitution of threonine 38 with aspartic acid. J. Biol. Chem. **272:** 2218–2222
31 Mazza C., Breton R., Housset D. and Fontecilla-Camps J. C. (1998) Unusual charge stabilization of $NADP^+$ in 17β-hydroxysteroid dehydrogenase. J. Biol. Chem. **273:** 8145–8152
32 Tsigelny I. and Baker M. E. (1995) Structures stabilizing the dimer interface on human 11β-hydroxysteroid dehydrogenase-types 1 and 2 and human 15-hydroxyprostaglandin dehydrogenase and their homologs. Biochem. Biophys. Res. Commun. **217:** 859–868
33 Benach J., Atrian S., Gonzalez-Duarte R. and Ladenstein R. (1998) The refined crystal structure of Drosophila lebanonensis alcohol dehydrogenase at 1.9 Å resolution. J. Mol. Biol. **282:** 383–399
34 Chenevert S. W., Fossett N. G., Chang S. H., Tsigelny I., Baker M. E. and Lee W. R. (1995) Amino acids important in enzyme activity and dimer stability for *Drosophila* alcohol dehydrogenase. Biochem. J. **308:** 419–423
35 Wermuth B. (1981) Purification and properties of an NADPH-dependent carbonyl reductase from human brain: relationship to prostaglandin 9-keto reductase and xenobiotic ketone reductase. J. Biol. Chem. **256:** 1206–1213
36 Tanaka M., Ohno S., Adachi S., Nakajin S., Shinoda M. and Nagahama Y. (1992) Pig testicular 20 beta-hydroxysteroid dehydrogenase exhibits carbonyl reductase-like structure and activity: cDNA cloning of pig testicular 20 beta-hydroxysteroid dehydrogenase. J. Biol. Chem. **267:** 13451–13455